AI Governance in the Financial Industry

Robin Feldman*

Kara Stein†

Abstract

Legal regimes in the United States generally conceptualize obligations as attaching along one of two pathways: through the entity or the individual. Although these dual conceptualizations made sense in an ordinary pre-modern world, they no longer capture the financial system landscape, now that artificial intelligence has entered the scene. Neither person nor entity, artificial intelligence is an activity or a capacity, something that mediates relations between individuals and entities. And whether we like it or not, artificial intelligence has already reshaped financial markets. From Robinhood, to the Flash Crash, to Twitter's Hash Crash, to the Knight Capital incident, each of these episodes foreshadows the potential for puzzling conundrums and serious disruptions.

Little space exists in current legal and regulatory regimes to properly manage the actions of artificial intelligence in the financial space. Artificial intelligence does not "have intent" and therefore cannot form the scienter required in many securities law contexts. It also defies the approach commonly used in financial regulation of focusing on size or sophistication. Moreover, the activity of artificial intelligence is too diffuse, distributed, and ephemeral to effectively govern by aiming regulatory firepower at the artificial intelligence itself or even at the entities and individuals currently targeted in securities law. Even when the law deviates from the classic focus on entities and individuals, as it meanders through areas that implicate artificial intelligence, we lack a unifying theory for what we are doing and why.

* Arthur J. Goldberg Distinguished Professor of Law, Albert Abramson '54 Distinguished Professor of Law Chair, Director, Center for Innovation (C4i), University of California Hastings Law.

† Board Member of the Public Company Accounting Oversight Board; Former Director of the AI & Capital Markets Initiative at the Center for Innovation, University of California, Hastings; Former Lecturer-in-Law, University of Pennsylvania Carey Law School; Former Commissioner, U.S. Securities and Exchange Commission (2013-2019). The authors would like to thank Jill E. Fisch, Doru Gavril, Joseph Grundfest, and Robert Peak for their ideas and comments on the draft, as well as Nathan Brown, Joseph Clateman, Maisam Goreish, Oriana Tang, David Toppelberg, and Todd Warshawsky for invaluable research assistance. We are particularly indebted to Gideon Schor for leading the research team. The views expressed in this paper are those of the authors only and do not necessarily reflect the views of the authors' employers or any other entities with which the authors might be associated.

To begin filling this void, we propose conceptualizing artificial intelligence as a type of skill or capacity—a superpower, if you will. Just as the power of flight opens new avenues for superheroes, so, too, does the power of artificial intelligence open new avenues for mere mortals. With the capacity of flight as its animating imagery, the article proposes what we would call "touchpoint regulation." Specifically, we set out three forms of scaffolding—touchpoints, types of evil, and types of players—that provide the essential structure for any body of law society will need for governing artificial intelligence in the financial industry.

Introduction

Legal regimes in the United States generally conceptualize obligations as attaching along one of two pathways: through the entity or through the individual. When a car accident occurs, tort law asks whether the driver or the auto manufacturer is at fault. When a partnership fails, bankruptcy law dictates whether the partnership and its assets will pay creditors, or whether the responsibility falls to the individual partners. Government regulation generally follows the same two pathways. Securities law regulates individual brokers but also regulates the investment bank at which the broker works. Even the notion of sovereign immunity—when and whether one can hold the government accountable—springs from the historic notion of the government in the form of the "sovereign," that is, the person who wears the crown.

In a pre-modern world, these conceptualizations make good sense. If society wishes to impose order on the chaos of human existence, where else would one focus beyond persons and groups of persons? Of course, the notion of entities such as

corporations or partnerships having legal life and identity did not exist from time immemorial.[1] Corporations were not always considered people. Nevertheless, the law has breathed life into corporations, partnerships, unions, PACs, fraternal organizations,[2] and the like, such that they now constitute targets of regulation.

The dual conceptualization of entity and individual as the essential targets of legal constraints no longer captures the landscape. Rather, artificial intelligence disrupts this entire notion. Artificial intelligence is not a person[3] or an entity. Nor does artificial intelligence respond to the rewards, punishments, and incentive structures that underlie much of the logic in legal doctrines—at least not standing alone. In fact, an artificial intelligence is not a "thing" at all. It is a process or an activity, a verb, something that mediates relations between individuals and entities. But how do we regulate a verb? Or perhaps more saliently, what does a regulatory paradigm for an activity look like?

The need for answers is fast approaching in our rearview mirror, threatening to overtake us at any moment. Whether we like it or not, artificial intelligence processes have already reshaped financial markets. Artificial intelligence already automates a broad array of trading activities, changes both decentralized market access and information, and increasingly directs human behavior towards more efficient outcomes. The role of artificial intelligence in financial markets, though still in its infancy, will only grow in the years to come, perhaps faster than we would like.[4]

Although the notion of artificial intelligence may sound mysterious and complex, its role in any market generally falls under a single definition. Artificial intelligence is any set of technologies capable of adaptive, predictive power in the context of solving a problem.[5] As scientific capacities stand now, the process normally involves using past

1. *See, e.g.*, Samuel Williston, *History of the Law of Business Corporations Before 1800* (pt. I), 2 HARV. L. REV. 105, 106 (1888) (outlining the ancient history of corporations). Since its inception in ancient Rome (or, some say, ancient Greece), the corporate form generally has been understood as a legal fiction or proxy created for practical benefit. Both prior and subsequent to the development of limited liability in England, jurists have recognized that corporations (as well as governments, partnerships, and even personified boats) are creatures of the law created for the purpose of attaching to them judicially recognizable powers and responsibilities. *See, e.g.*, Trs. of Dartmouth Coll. v. Woodward, 17 U.S. 518, 636 (1819) ("A corporation is an artificial being, invisible, intangible, and existing only in contemplation of law. Being the mere creature of law, it possesses only those properties which the charter of its creation confers upon it, either expressly, or as incidental to its very existence.").
2. *See, e.g.*, I.R.C. § 501(c)(8) (federal income tax code exemption for fraternal beneficiary societies); I.R.C. § 501(c)(10) (federal income tax code exemption for domestic fraternal societies).
3. For an extensive discussion of the limitations of personhood as a legal approach for conceptualizing AI, see Nadia Banteka, *Artificially Intelligent Persons*, 58 HOUS. L. REV. 537 (2021).
4. *See, e.g.*, Longbing Cao, *AI in Finance: A Review*, ACM COMPUT. SURV., Mar. 2018, at 1 (2018) (discussing trends in artificial intelligence's use in finance).
5. *See* Kara M. Stein, Former Comm'r of the U.S. Sec. & Exch. Comm'n, Keynote Address for the Abraham L. Pomerantz Lecture at Brooklyn Law School: Investor Protection in the Digital Age (Sep. 24, 2019), *in* 85 BROOK. L. REV. 631, 635 (2020).

data to train a model to make predictions on future data,[6] and then *improving* those predictions as new data enters the model. The most important aspect to understand about artificial intelligence, however, is the following: In contrast to earlier predictive analytic methods, artificial intelligence can make assumptions, test, learn, reiterate, and improve.

Defining artificial intelligence may be relatively simple, but understanding what it *does* is considerably more difficult. Frequently referred to as a "black box," many artificial intelligence processes remain mysterious to outside observers. While the original artificial intelligence code may have been written by human programmers,[7] artificial intelligence learns by utilizing thousands upon thousands of data points in an endless stream to retrain its internal models. For good reason, no human mind could possibly keep up, and even unwinding the pathways taken can be challenging—at least in any way that will be readily understandable to humans.

This leap in analytical power also makes artificial intelligence potentially dangerous for any market it operates within. Artificial intelligence guides human behavior—sometimes without those very humans being aware of the implications—and may distort market structures, compound the advantages held by the biggest actors, promote quasi-collaborative and anticompetitive behaviors, and, if unaccounted for, undermine legal and financial certainty.[8]

Little space exists in current legal and regulatory regimes to properly manage the

6. Rather than simply improving based on training data fed to it, artificially intelligent programs may soon be able to proactively take in data from a new environment, effectively "training" itself. *See* Craig S. Smith, *Computers Already Learn From Us. But Can They Teach Themselves?*, N.Y. TIMES (Jul. 16, 2021), https://perma.cc/RQ7T-MYZ8 (describing the development of self-supervised learning as the next frontier in AI); H. James Wilson, Paul R. Daugherty & Chase Davenport, *The Future of AI Will Be About Less Data, Not More*, HARV. BUS. REV. (Jan. 14, 2019), https://perma.cc/WAH7-E9LU (describing Gaussian processes, exemplified by Google's Project Loon, that combine past and real-time data to make predictions based on probability). Machine learning is already used to create and improve the data sets needed to train other AI systems. *See, e.g.*, Connor Shorten & Taghi M. Khoshgoftaar, *A Survey on Image Data Augmentation for Deep Learning*, J. BIG DATA, 2019, at 1 (describing the use of algorithms to augment training datasets); Ron Miller, *Superb AI Generates Customized Training Data for Machine Learning Projects*, TECH CRUNCH (Feb. 25, 2019, 12:12 PM PST), https://perma.cc/G2KR-UNVP (describing a company that uses artificial intelligence to customize and develop training data for other AI projects).

7. Under development are AI programs that are able to compose new programs, so that the role of humans would be limited to defining the desired output or purpose of the AI-authored algorithm. *See* Will Knight, *Now For AI's Latest Trick: Writing Computer Code*, WIRED (Apr. 23, 2021, 7:00 AM), https://perma.cc/4ACN-AVLH (describing GPT-3, an artificial neural network, which could lead to tools that generate code based on a text description of what the program should accomplish); Dom Galeon, *Our Computers Are Learning How To Code Themselves: Human Coders Beware*, FUTURISM (Feb. 25, 2017), https://perma.cc/9MUW-S4JK (describing machine learning programs that employ program synthesis to compile existing lines of code into a program with a targeted output).

8. *See generally* Alessio Azzutti, Wolf-Georg Ringe & H. Siegfried Stiehl, *Machine Learning, Market Manipulation and Collusion on Capital Markets: Why the 'Black Box' Matters* (Eur. Banking Inst., Working Paper Series No. 84, 2021), https://ssrn.com/abstract=3788872.

actions of artificial intelligence in the financial arena.[9] Artificial intelligence does not "have intent" and therefore cannot form the scienter required in many securities law contexts. Of course, an abstraction such as an entity cannot hold any "intent," either. In the case of entity regulation, however, one can extrapolate through the intent or actions of the individuals who hold positions of authority and make decisions on behalf of the entity. A corporation may need counsel separate from its officers and directors if their interests diverge, but when the law looks for scienter on the corporation's behalf, it is looking at the scienter of the human actors.

Even with the rudimentary stages of artificial intelligence that exist today, however, the "intent" of artificial intelligence is not the same as the intent of any particular human. Through its updating and self-development, artificial intelligence distances itself from its creator, posing further accountability and regulatory issues. A programmer may be found guilty of encoding deliberate malice into a system, but many cases will require more subtle discernment.[10] Increasingly capable software will continue to challenge intuitive understandings of responsibility and liability, as algorithms "learn" beyond their initial coding or design. Consequently, the autonomy of artificial intelligence offers opportunities for plausible deniability. What is the recourse when a trader blames the artificial intelligence? Who is at fault when an artificial intelligence directs corporate activity in a way that violates the law or imposes risk of harm to humans? How should enforcement agencies react when an entire group of people moves in concert because the moderating artificial intelligence directed it? These are the types of challenging issues that law related to financial markets and institutions is entirely unprepared to address. And of course, it goes without saying that artificial intelligence has the potential to greatly benefit society, rather than simply imposing

9. Currently, AI-specific regulation is sparse. The FTC has issued guidance on the use of AI, but it consists mainly of explaining how the use of artificial intelligence generally falls within the scope of older laws, such as the Fair Credit Reporting Act (FCRA) and Equal Credit Opportunity Act (ECOA). *See* Andrew Smith, *Using Artificial Intelligence and Algorithms*, FED. TRADE COMM'N: BUS. BLOG (Apr. 8, 2020, 9:58 AM), https://perma.cc/S54G-WY8G. The SEC's guidance to this point is limited, and a recent FINRA report shows how artificial intelligence's use currently outpaces regulation specifically addressing it. *See* FIN. INDUS. REG. AUTH., ARTIFICIAL INTELLIGENCE (AI) IN THE SECURITIES INDUSTRY 1, 11-19 (2020). Although the Office of Management and Budget recently finalized guidance to federal agencies on the regulation of artificial intelligence technologies, the most comprehensive regulation of artificial intelligence remains the 2019 National Defense Authorization Act § 1051, which established the National Security Commission on Artificial Intelligence to review and compose reports proposing controls for the use of artificial intelligence in the security and defense sector. RUSSELL T. VOUGHT, OFF. OF MGMT. & BUDGET, EXEC. OFF. OF THE PRESIDENT, MEMORANDUM FOR THE HEADS OF EXECUTIVE DEPARTMENTS AND AGENCIES 1-2 (2020); John S. McCain National Defense Authorization Act for Fiscal Year 2019, Pub. L. No. 115-232, § 1051, 132 Stat. 1636, 1962-65 (2018).

10. *See, e.g.*, Iria Giuffrida, *Liability for AI Decision-Making: Some Legal and Ethical Considerations*, 88 FORDHAM L. REV. 439, 446-447 (2019) (offering real and hypothetical examples of the difficulty of determining causation when artificial intelligence is involved in a tort context); *see also* Yavar Bathaee, *Artificial Intelligence Opinion Liability*, 35 BERKELEY TECH. L.J. 113 (2020) (discussing the problems of trying to regulate artificial intelligence through intent-based heuristics); Azzutti et al., *supra* note 8, at 29-31.

the risk of harm. One would not want to operate in Luddite fashion and imagine that the march of science could be turned back at the wave of a hand.

Artificial intelligence not only defies categorization as entity or individual; it also creates different relationships between individuals and entities. For example, artificial intelligence has facilitated the creation of a new species of spaces that press against the logic of entity as a locus of legal rules. These not only include the creation of dark web financial interactions and exchanges; they also include a web of time and space in which interactions can occur in minuscule fractions of a second when different automated-trading programs alter their trajectories as they bounce off each other across different exchanges.[11] Artificial intelligence takes these issues to new depths. Quite simply, in these dark and distributed domains, the concept of entity loses both its descriptive power and its regulatory force. Thus, attempts to govern artificial intelligence in these areas will inevitably be stymied by the fact that artificial intelligence cannot fit within the existing categorizations of the governed under the law. From a theoretical standpoint, we have no one to hold accountable. From a practical standpoint, we have nothing to grab ahold of, anyway.

Grounding regulation in the concept of entity or individual distorts and limits the way we might think about the legal system's potential embrace of artificial intelligence. It isn't that everything about individual and entity is wrong in the context of artificial intelligence; the problem is that the resulting picture is out of focus—and at times, downright useless. In effect, we lack a unifying theory for what we are doing and why.

Although society cannot wrap its arms around artificial intelligence in the form of a person or entity, not all is lost. We are not doomed to allow artificial intelligence to run rampant across the landscape. This article proposes conceptualizing artificial intelligence as a type of skill or capacity—a superpower, if you will. Just as the power of flight opens new avenues for superheroes, so, too, does the power of artificial intelligence open new avenues for mere mortals. This article explores how the concept of flight and the activity of flying provide a loose analogy for thinking about the regulation of artificial intelligence in the financial industry.

With the capacity of flight as its animating imagery, this article proposes what we would call "touchpoint regulation." Specifically, we set out three forms of scaffolding that are designed to support a regulatory framework for artificial intelligence. The key,

---

11. *See* Yesha Yadav, *The Failure of Liability in Modern Markets*, 102 VA. L. REV. 1031, 1079-81 (2016) (describing the interconnectedness of automated high-frequency trading algorithms, and the systemic ripple effects that can result across financial markets); *see also* Matthew U. Scherer, *Regulating Artificial Intelligence Systems: Risks, Challenges, Competencies, and Strategies*, 29 HARV. J.L. & TECH. 353, 365-66, 372-73 (2016) (describing the range of expected and unexpected behavior artificial intelligence could manifest, with or without a designer's intent); Jack B. Balkin, *The Path of Robotics Law*, 6 CALIF. L. REV. CIR. 45, 52 (2015) ("But although the risk of *some* kind of injury at *some point* in the future is foreseeable whenever one introduces a new technology, how and when an injury occurs may not be particularly foreseeable to . . . potential defendants . . . . If the law hopes to assign responsibility to humans and corporations, injuries by robotic and AI systems may strain traditional concepts of foreseeability.").

initial concept flows from the notion that the activity of artificial intelligence is too diffuse, distributed, and ephemeral to effectively govern by aiming regulatory firepower at the artificial intelligence itself or even at the entities and individuals currently targeted in securities law. Instead, the law should regulate through the various touchpoints at which artificial intelligence activities connect to the broader financial system and through which one could hold all actors—not just certain entities or licensed professionals—accountable.

Second, the framework organizes remedies and culpability according to the following categories of harms: programmed harms, reasonably predictable harms, and entirely unpredictable harms.[12] In other words, in determining responsibility and remedy when bad things occur, the law should ask whether it is the evil you planned, the evil you could have predicted, or the evil you couldn't reasonably have expected.[13] These approaches leave room for the potentially unknowable aspects of artificial intelligence without permitting wholesale abdication of responsibility by allowing a party to say, "The algorithm made me do it."

Finally, the touchpoint framework organizes the obligations of touchpoint actors according to whether they are users, intermediaries and gatekeepers, or creators. When actors reach the touchpoints—connecting with the broader financial systems through which we will hold all actors accountable—their actions will be examined based on the role they play in the system. When the type of actor cannot be determined at a touchpoint, the law can assign the highest level of responsibility, thereby creating incentives for parties to identify themselves, rather than hide in the shadows.

These three scaffolding structures—touchpoints, types of evil, and types of players—provide the essential structure for any body of law that society will need to govern artificial intelligence in the financial industry. Part I of this article describes the current capabilities of artificial intelligence and the dependence of existing financial regulatory regimes on intent in determining liability. Part II identifies a variety of problems artificial intelligence may pose to financial regulatory regimes, identifying the potential for significant disruptions and distortions. As red flags signifying the danger ahead, this part takes as examples a combination of hypothetical situations and recent financial episodes—such as Robinhood, the Flash Crash, and Twitter's Hash Crash—that foreshadow the potential for puzzling conundrums and serious disruptions. Part III poses the analogy of flight as a method of conceptualizing how to govern artificial intelligence in the financial sector. As with artificial intelligence, air travel can lead to good results or bad, is subject to error (both human and mechanical, as well as unpredictable disasters), and occurs at a precise moment of activity involving the intersection of various parties whose interests and activities must be intricately managed. Building on this concept, Part III also defines the concept of touchpoint regulation, which is grounded in the notion that certain aspects of U.S. financial markets

12. *See* Robin Feldman, Professor of Law, Emerging Competition, Innovation, and Market Structure Questions Around Algorithms, Artificial Intelligence, and Predictive Analytics at the FTC Hearings on Competition and Consumer Protection in the 21st Century: FTC Hearing No. 7 106 (Nov. 14, 2018) (transcript available at https://perma.cc/GZH6-KYRG).

13. *See id.*

operate as desirable entry points, and therefore desirable avenues for regulatory control. Finally, this Part suggests organizing such regulation according to the predictability of the bad outcome along with the types of actors involved in the activity—users, creators, or intermediaries. Together, this Article sets out a paradigm for moving forward with effective regulation.

I. Understanding Artificial Intelligence in the Financial Industry

Understanding the collision course on which artificial intelligence and the financial markets are set requires that one understand the capacities and limitations of artificial intelligence as well as the baseline assumptions built into existing financial regulatory systems.

A. What Is Artificial Intelligence?

Artificial intelligence is not just a computer algorithm for predictive analytics. That is a relatively simple beast. In its broadest sense, a computer algorithm is merely any series of steps that a computer forms on input data. For some time, one has been able to develop such series of steps for the purpose of making predictions. For example, researchers in many fields have been able to create mathematical models, implemented on computers, that offer predictions of everything from what a consumer might buy, to whether a building will withstand an earthquake, to whether an adversary will go to war.

In contrast, when one speaks about artificial intelligence, most people are referring to deep learning. Given the current state of artificial intelligence, deep learning means using past data to train a model that can, on its own, make predictions on future data and direct choices based on those predictions.[14] (For example, will that pedestrian step in front of the car, and should the car apply the brakes?) Most importantly, artificial intelligence—again, on its own—can develop assumptions, test those assumptions, learn, and reiterate in ever-increasing layers of analysis.

Deep learning is an especially prominent and fashionable sector of artificial intelligence.[15] Although an extensive primer would be out of scope, deep learning, in brief, is a catch-all term that refers to a type of mathematical model used to analyze data. Drawing on an analogy of how the human brain processes information, these so-called "neural networks" analyze data by administering a sequence of mathematical transformations. Each "layer" of the neural network is comprised of an additional transformation; "deep learning" describes the study of neural networks with many such

14. Machine learning programs show promise to advance beyond this capacity to effectively be able to design their own programs. *See* Knight, *supra* note 7 (describing a neural network that can be used to generate code based on a text description of what the program should accomplish); Galeon, *supra* note 7 (describing machine learning programs that can compile existing lines of code to design a program that completes a discrete task).

15. For more in-depth discussion of these AI concepts and their implications for the legal system, see Robin Feldman, *Artificial Intelligence: The Importance of Trust & Distrust*, 21 GREEN BAG 201 (2018).

layers. "Training" the neural network—as the iterative process that activates the network's many layers is termed—requires "training data." The efficacy of any deep learning project, as a result, relies on robust and reliable training data.

Artificial intelligence has progressed at an astounding clip. Deep learning, which serves as the foundation of the entire discipline, only became viable after a 2006 paper outlined how to train neural networks efficiently.[16] The generative adversarial models that enable most neural networks today did not develop until 2014.[17] Staggering advances in the industry have been obtained since, comparable to what would have required lifetimes in other fields.

Despite these tremendous advances, the artificial intelligence programs we use today do not yet approach the comprehensive general intelligence of a human mind. All current artificial intelligence is "weak," a description that recognizes that artificial intelligence is limited to completing tasks based on the data it is given.[18] "Strong" artificial intelligence, by contrast, can be ascribed something like a consciousness in its performance of tasks—an ability that some scientists and philosophers refuse to even concede as a possibility.[19] A helpful Hollywood analogue to understand the difference between weak and strong forms of artificial intelligence can be found in comparing Iron Man to the Terminator.[20] The former enhances the abilities of the human who controls it; the latter acts autonomously.[21]

In fact, the human-machine synthesis of Iron Man overstates the scope of present-day weak artificial intelligence, whose discrete applications make it more like a tool. The "narrow" artificial intelligence programs that predominate are coded to complete a specific, bounded task, such as recommending television programs or filtering spam.[22] Although narrow artificial intelligence programs have advanced to proliferate across a range of contexts, the prospect of a "general" artificial intelligence—a single program able to accomplish any human task, short of requiring consciousness—remains untenable.[23] Artificial intelligence, that is, should not yet be understood as a

16. Geoffrey E. Hinton, Simon Osindero & Yee-Whye Teh, *A Fast Learning Algorithm for Deep Belief Nets*, 18 NEURAL COMPUTATION 1527 (2006).

17. *See* Ian J. Goodfellow Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville & Yoshua Bengio, *Generative Adversarial Nets*, NEURAL INFO. PROCESSING SYS. PROC., 2014, at 1.

18. *See* Feldman, *supra* note 15, at 202; *see also* Bernard Marr, *What Is Weak (Narrow) AI? Here Are 8 Practical Examples*, BERNARD MARR & CO., https://perma.cc/GM6B-D7AT (archived Nov. 7, 2021) (giving examples of weak AI, such as Apple's voice-activated assistant, Siri).

19. See Feldman, *supra* note 15, at 202 for more on the distinctions between weak and strong artificial intelligence.

20. *See id.* at 204.

21. *See id.*

22. *See* Marr, *supra* note 18; *Narrow AI*, DEEPAI, https://perma.cc/MVC5-3HHF (archived Nov. 7, 2021) (describing and exemplifying narrow AI).

23. General artificial intelligence, lacking consciousness, still constitutes weak artificial intelligence, although it is much closer to strong artificial intelligence in its versatility between tasks. *See* Ragnar Fjelland, *Why General Artificial Intelligence Will Not Be Realized*, HUMANITS. SOC. SCI. COMMUN., 2020, at 1, 2; Feldman, *supra* note 15, at 202. Others have

means of replacing human intelligence. Instead, modern artificial intelligence programs, as with tools, function to expand the human capacity to complete a task.

The financial industry is increasingly familiar with the enhancing ability of artificial intelligence. In addition to trading stocks, companies have deployed artificial intelligence to assess loan applicants,[24] predict credit card fraud,[25] underwrite insurance policies,[26] customize individual financial planning[27] and commercial banking programs,[28] and automate regulatory compliance.[29] Moreover, the data-intensive nature of finance and banking—not to mention the diligent record keeping financial regulations mandate—makes the sector particularly receptive to future machine learning forays.

The increased use of artificial intelligence, however, may not imply increased transparency or legibility. Rather, the narrow scope, and concomitant lack of "common sense" characteristic of today's artificial intelligence, manifest in unexpected and often inexplicable outcomes. For instance, an algorithm designed to classify street signs misread stop signs as speed limits when stickers were pasted to the sign;[30] the

posited that general artificial intelligence should be understood as strong artificial intelligence; in either case, general artificial intelligence has yet to be realized—and may be as much a pipe dream as the notion of strong, conscious artificial intelligence. *See* Ben Dickson, *What Is Narrow, General And Super Artificial Intelligence*, TECHTALKS (May 12, 2017), https://perma.cc/FGW7-P8YJ. Still other commentators have insisted that the staggering advances in processing power will lead to so-called "super" artificial intelligence forms, surpassing general human intelligence in the near future. *See, e.g.*, Nick Bostrum, *How Long Before Superintelligence?*, 5 LINGUISTIC & PHIL. INVESTIGATIONS 11 (2006).

24. *See 5 Ways AI Is Transforming the Finance Industry*, MARUTI TECHLABS, (archived Nov. 7, 2021) (describing the role of artificial intelligence in applicant risk assessment). The introduction of artificial intelligence to the process of issuing credit has raised questions about whether machine learning ameliorates or exacerbates historically racist lending practices. *See also infra* text accompanying notes 37-38; *cf.* Anya E.R. Prince & Daniel Schwarcz, *Proxy Discrimination in the Age of Artificial Intelligence and Big Data*, 105 IOWA L. REV. 1257, 1264, 1288 (2020) (arguing that artificial intelligence poses problems for anti-discrimination regimes by using suspect proxies to make decisions in areas such as credit lending). *Compare* Robert P. Bartlett, Adair Morse, Richard Stanton & Nancy Wallace, *Consumer-Lending Discrimination in the FinTech Era* 4 (Nat'l Bureau of Econ. Rsch., Working Paper No. 25943, 2019) (analyzing algorithmic lending and finding that artificial intelligence tended to issue higher interest rates to protected minorities), *with* Sian Townson, *AI Can Make Bank Loans More Fair*, HARV. BUS. REV. (Nov. 6, 2020), https://perma.cc/F5TH-X7FK (outlining how artificial intelligence can rectify historically racist credit decisions).

25. Alyssa Schroer, *AI and the Bottom Line: 20 Examples of Artificial Intelligence in Finance*, BUILT IN (Aug. 6, 2021), https://perma.cc/TLC9-GFYQ.

26. *See* FIN. SERVS. INST., PWC, TOP FINANCIAL SERVICES ISSUES OF 2018 7 (2018), https://perma.cc/G99C-ZUNT.

27. *See* Jay Adkisson, *Artificial Intelligence Will Replace Your Financial Adviser—And That's a Good Thing*, FORBES (Jan. 23, 2019, 2:17 PM EST), https://perma.cc/NLY7-F4WK.

28. *See* Arthur Bachinskiy, *The Growing Impact of AI in Financial Services: Six Examples*, TOWARDS DATA SCIENCE (Feb. 21, 2019), https://perma.cc/6NMD-HQFT (describing personalized commercial banking).

29. *See id.* (describing how Robotic Process Automation (RPA) allows banks to automatically comply with data collection regulations such as "Know Your Customer").

30. *See* Jonathan M. Gitlin, *Hacking Street Signs with Stickers Could Confuse Self-Driving Cars*, ARS TECHNICA (Sept. 1, 2017, 9:30 AM), https://perma.cc/T9LG-RBYT.

consequences for a self-driving car equipped with this technology are not difficult to imagine. Similar problems have plagued the deep learning networks that operate chatbots and facial recognition software, too, with one artificially intelligent camera mistaking a referee's bald head for the soccer ball it was programmed to track.[31]

One cannot overemphasize the extent to which artificial intelligence systems behave in ways that are entirely unpredictable. For example, Free University of Berlin computer science professor Raul Rojas built a smart house in 2009 in which everything (lights, music, television, heating and cooling, etc.) was connected to the Internet.[32] In 2013, the entire house froze and stopped responding to Rojas's commands. The professor discovered that a light bulb had burned out; in its efforts to alert "the hub," the bulb sent out continuous requests that overloaded the network, also known as a "denial of service attack." Rojas emphasized that "the bulb receptacle [was] not supposed to do this," and that the result would have been "horrible" for someone living in the house who was not a computer science professor.[33] "They would have torn down the wiring in the house trying to figure out what was wrong."[34]

Understanding the pathways that led an artificial intelligence algorithm to act in a particular manner can be challenging, even for those who designed the program. This so-called "black box" nature of artificial intelligence plays a prominent role in debates over the emergence of lethal autonomous weapons systems (LAWS).[35] Of particular note is the concern that "humans may no longer be able to predict who or what is made the target of an attack, or even explain why a particular target was chosen by a LAWS."[36]

Even without malfunctioning, some algorithms serve to entrench significant material disparities, raising calls for their correction. In stark contrast to the vision of an unbiased artificial intelligence, algorithms frequently amplify existing discriminatory practices in finance, policing, and health care.[37] Utilized in commercial banking, for

31. For a more in-depth explanation of this example and other similarly headline-grabbing blunders, see *2020 in Review: 10 AI Failures*, SYNCED (Jan. 1, 2021), https://perma.cc/U2QT-6GG4.

32. Kashmir Hill, *This Guy's Light Bulb Performed a DoS Attack on His Entire Smart House*, SPLINTER (Mar. 3, 2015, 9:41 AM), https://perma.cc/D3VD-9EKL.

33. *Id.*

34. *Id.*

35. *See, e.g.*, REGINA SURBER, ICT4PEACE FOUND., ARTIFICIAL INTELLIGENCE: AUTONOMOUS TECHNOLOGY (AT), LETHAL AUTONOMOUS WEAPONS SYSTEMS (LAWS), & PEACE TIME THREATS 10 (2018), https://perma.cc/YF8U-F7PK; Aiden Warren & Alek Hillas, *When Robots Go to War: The Future of Unmanned Conflict*, YALE J. INT'L. AFF. (May 17, 2017), https://perma.cc/LC2G-9DDK.

36. *See* SURBER, *supra* note 35, at 10; *cf.* Peter M. Asaro, *AI Ethics in Predictive Policing: From Models of Threat to an Ethics of Care*, IEEE TECH. & SOC'Y MAG., June 2019, at 40, 41 (describing similar concerns in the context of predictive policing).

37. *See, e.g.*, SIMONE BROWNE, DARK MATTERS: ON THE SURVEILLANCE OF BLACKNESS (2015); Mutale Nkonde, *Automated Anti-Blackness: Facial Recognition in Brooklyn, New York*, 2019-20 HARV. KENNEDY SCH. J. AFR. AM. POL'Y 30, 32; Alexander Campolo & Kate Crawford, *Enchanted Determinism: Power Without Responsibility in Artificial Intelligence*, ENGAGING SCI., TECH. & SOC'Y., 2020, at 1, 5; Aurore Lentz, *Garbage In, Garbage Out: Is AI Discriminatory or*

instance, machine learning generates comparatively higher interest rates for minorities applying for loans.[38] Diagnostic and patient care algorithms return similarly prejudicial results, such that Black patients are far less likely to be referred for extra medical care than their white counterparts.[39] The unintended biases of artificial intelligence should, at least for the present, only reinforce the subordination of the technology to the humans and entities that harness it.

In brief, artificial intelligence is not a person or an entity. It is an activity, a capacity, or tool that can be added to enhance human skill. For the current and foreseeable future, the role of artificial intelligence will continue as a form of human enhancement. Moreover, the learning and reiterative nature of artificial intelligence means that the processes dictated by artificial intelligence may move in directions independent of, or unanticipated by, human users and creators.

B. Whose Mind Is It Anyway?[40]

Although artificial intelligence may provide a means for human enhancement, it lacks any human incentives, motivations, or intent—at least for the foreseeable future.[41] It is easy to speak of an artificial intelligence as "choosing" or "behaving" or even "cheating,"[42] but characterizations of this kind are simply clever anthropomorphisms, attributing human characteristics to an inanimate process. The operation of the machine learning program contains no human thought, motivation, or intent—beyond, of course, the intent of the developer. And yet, when the self-learning capacity of the artificial intelligence leads it to advance in ways unintended and unforeseen by the

---

Simply a Mirror of IRL Inequalities?, UNIVERSAL RIGHTS GROUP (Jan. 18, 2021), https://perma.cc/V2BH-WRVJ; FREDERIK ZUIDERVEEN BORGESIUS, DISCRIMINATION, ARTIFICIAL INTELLIGENCE, AND ALGORITHMIC DECISION-MAKING (2018), https://perma.cc/8RAC-56C2; *see also* discussion *supra* note 24.

38. *See* Bartlett et al., *supra* note 24.

39. *See* Ziad Obermeyer, Brian Powers, Christine Vogeli & Sendhil Mullainathan, *Dissecting Racial Bias in an Algorithm Used to Manage the Health of Populations*, 366 SCIENCE 447, 449 (2019) (finding that addressing racial disparity in a prominent diagnostic algorithm would increase by almost half the percentage of Black patients who receive additional care). *See generally* Sharona Hoffmann & Andy Podgurski, *Artificial Intelligence and Discrimination in Health Care*, YALE J. HEALTH POL'Y, L. & ETHICS, 2020, at 1, 17-19 (2020) (describing other examples of discriminatory healthcare algorithms, including a diagnostic algorithm that may misdiagnose benign vs. malignant tumors in dark-skinned patients; an Alzheimer's diagnostic algorithm that misdiagnoses non-native English speakers; a patient-scheduling program that was more likely to double-book poorer patients).

40. For the title of this subsection, I pay homage to the comedy improv television show *Whose Line Is It Anyway?* (ABC television broadcast).

41. *But cf.* Ron Dolin, *Defining the Problem—Regulating Behavior*, THINK OUTSIDE THE BAR (Jan. 8, 2020), https://perma.cc/TKE5-RFDZ (raising the tantalizing possibility, in the context of a hypothetical of the "naughty vacuum cleaner," that one can apply concepts of accountability, responsibility, and deterrence to machine learning).

42. *See infra* text accompanying notes 95-97 (describing an AI that cheated on a task by hiding information from its creator that could later be accessed to accomplish the task more easily).

developer, it may be difficult to conclude that the developer is at fault. Moreover, the law may find it difficult to pinpoint an individual developer when the interconnectivity of different algorithms makes it difficult to determine where one algorithm ends and the other begins. Modern judicial and regulatory regimes, however, rely on concepts of motivation and intent to assess misconduct and assign liability.

For example, anti-fraud measures, such as the Securities and Exchange Commission (SEC)'s Rule 10b-5, are designed to deter securities fraud. Doing so usually depends on demonstrating that a defendant acted with scienter,[43] which refers to a mental state characterized by an intent to deceive or defraud.[44] On one hand, Section 10(b) of the 1934 Exchange Act, and the SEC's corresponding Rule 10b-5, have proven effective for over eighty-five years due, in part, to the broad and general language used that forbids fraudulent behavior in the securities market. Specifically, the rule prohibits any person from employing "any device, scheme, or artifice to defraud"[45] or engaging in any deceitful business practice linked to the purchase or sale of a security.[46] Whatever may be the outer bounds of this capacious wording, artificial intelligence fits comfortably within it. As a *tool* for human advancement, artificial intelligence is clearly a "*device*." And, indisputably, *artificial* intelligence is an "*artifice*." Thus, although the rule, promulgated by the SEC in 1942, hardly could have foreseen the development of trading algorithms, its "catch-all" language aptly accommodates artificial intelligence as a potential tool of fraud.[47]

Subsequent court decisions and congressional amendments, however, have limited the reach of the rule by tying it firmly to scienter.[48] For example, in *Ernst & Ernst v.*

43. *Cf. supra* text accompanying notes 41-42.

44. *Scienter*, MERRIAM-WEBSTER.COM, https://perma.cc/D5PX-7J3P (archived Apr. 15, 2021).

45. 17 C.F.R. § 240.10b-5(a) (2020).

46. 17 C.F.R. § 240.10b-5(c) (2020).

47. *See* Eric C. Chaffee, *Standing Under Section 10(b) and Rule 10b-5: The Continued Validity of the Forced Seller Exception to the Purchaser-Seller Requirement*, 11 U. PA. J. BUS. L. 843, 843-844 (2009) (outlining legislative history of Section 10(b) and the subsequent codification in Rule 10b-5).

48. *See* Daniel A. McLaughlin & Mark Taticchi, *Corporate Scienter Under Section 10(b) and Rule 10b-5*, 46 SEC. REG. & L. REPORT 875 (2014), https://perma.cc/5744-86CJ ("Despite the lack of consensus on how to describe the rule or even on which courts have adopted it, however, a review of Section 10(b) cases reveals that, in practice, the courts have adopted a *de facto* rule grounded in traditional principles of agency law as set forth in the Restatement of Agency: *a corporation can violate Section 10(b) only when at least one of its employees or authorized agents knowingly or recklessly violates Section 10(b) in the scope of his or her employment.*"); Nicole M. Briski, Comment, *Pleading Scienter Under the Private Securities Litigation Reform Act of 1995: Did Congress Eliminate Recklessness, Motive, and Opportunity?*, 32 LOY. U. CHI. L.J. 155, 157 (2000) ("Every circuit to address the issue has agreed that the PSLRA [Private Securities Litigation Reform Act of 1995] has heightened the pleading standard for scienter that a plaintiff must meet in order to withstand a motion to dismiss. The circuits, however, disagree on the appropriate standard to adopt for pleading scienter under the PSLRA. Specifically, the circuits are split as to whether scienter under the PSLRA encompasses recklessness or, rather, a deliberate and knowing state of mind."); Ernst & Ernst v. Hochfelder, 425 U.S. 185, 193 (1976); Private Securities Litigation Reform Act of 1995, Pub. L. No. 104-67, 109 Stat. 737 (codified as amended in scattered sections of

Hochfelder, the Supreme Court indicated that alleging a defendant's "intent to deceive, manipulate or defraud" is a prerequisite to proving any Rule 10b-5 violation.[49] The case also clarified that merely showing that a defendant acted negligently does not suffice under this rule.[50]

In addition, the 1995 Private Securities Litigation Reform Act (PSLRA)[51] was passed in order to combat what was seen as a flood of frivolous securities lawsuits against corporations by private plaintiffs, as opposed to the SEC or the DOJ.[52] The law increased the difficulty of proving a Rule 10b-5 case, requiring private plaintiffs to demonstrate facts "giving rise to a strong inference" that a defendant acted with scienter.[53] As the Supreme Court decided a decade later in *Tellabs, Inc. v. Makor Issues & Rights, Ltd.*, an inference of scienter, to be strong, "must be cogent and at least as

15 U.S.C.).

49. *See Ernst & Ernst*, 425 U.S. at 193 ("We granted certiorari to resolve the question whether a private cause of action for damages will lie under § 10(b) and Rule 10b-5 in the absence of any allegation of 'scienter'—intent to deceive, manipulate, or defraud. We conclude that it will not and therefore we reverse." (citation omitted)).

50. *See id.* at 210 ("We think these procedural limitations indicate that the judicially created private damages remedy under § 10(b)—which has no comparable restrictions—cannot be extended, consistently with the intent of Congress, to actions premised on negligent wrongdoing." (footnote omitted)). However, *Sundstrand Corp. v. Sun Chemical Corp.* soon after affirmed that a reckless omission of information may be sufficient to establish scienter under Rule 10b-5. *See* Sundstrand Corp. v. Sun Chem. Corp., 553 F.2d 1033, 1044 (7th Cir. 1977) ("[W]e hold that a reckless omission of material facts upon which the plaintiff put justifiable reliance in connection with a sale or purchase of securities is actionable under Section 10(b) as fleshed out by Rule 10b-5."). Some circuits have upgraded this "recklessness" standard to "severe" or "deliberate" recklessness, significantly limiting the circumstances in which Rule 10b-5 can be enforced. *See* Julie Ann Sullivan, *The Science of Scienter: The Private Securities Litigation Act's Effect and the Long-Awaited Decision of* Makor Issue & Rights, Ltd. v. Tellabs, Inc., 1 SEVENTH CIR. REV. 327, 352-354 (2006) (noting that before 1995 recklessness was sufficient to establish scienter, that the 1995 passage of the Private Securities Litigation Reform Act did not change the substantive standard of scienter, but that, according to subsequent holdings of the Ninth and Eleventh Circuits, only "severe" or "deliberate" recklessness is sufficient to establish scienter).

51. Private Securities Litigation Reform Act of 1995.

52. *See Private Securities Litigation Reform Act Law and Legal Definition*, USLEGAL.COM, https://perma.cc/Z3FZ-BDGD (archived Nov. 6, 2021). The heightened pleading standards laid out by the PSLRA do not apply to enforcement actions brought by the SEC or DOJ. *See, e.g.*, U.S. SEC v. ICN Pharmaceuticals, Inc., 84 F. Supp. 2d 1097, 1099 (C.D. Cal. 2000) ("Defendant incorrectly argues that the heightened standard of the Private Securities Litigation Reform Act ('hereinafter 'PSLRA') of 1995 should be applied and that the SEC therefore has failed to 'plead specific facts creating a strong inference of scienter.' However, the 'more rigorous' pleading requirements under the PSLRA . . . only apply to private securities fraud actions . . . ."); *see also* Brief for the United States as Amicus Curiae Supporting Petitioners at 1, Tellabs, Inc. v. Makor Issues & Rts., Ltd., 551 U.S. 308 (2007) (No. 06-484) ("The issue in this case concerns the interpretation of the heightened pleading requirement for state of mind in *private securities fraud cases*. Meritorious private actions are an essential supplement to criminal prosecutions and civil enforcement actions brought, respectively, by DOJ and the SEC." (emphasis added)).

53. 15 U.S.C. § 78u-4(b)(2)(A).

compelling as any opposing inference of nonfraudulent intent."[54] Prior to passage of the PSLRA, some courts had been using the reasonable person standard, which did not require considering any competing interests in the defendant's favor. Thus, the *Tellabs* analysis of the PSLRA added to the trend of elevating the bar for demonstrating scienter in securities law.[55]

In addition to federal law, some state corporate governance laws[56] would fit poorly with holding directors liable for corporate activity involving artificial intelligence. Under Delaware law,[57] directors have two fiduciary duties: the duty of loyalty, which requires directors to put the corporation's best interests above any other interest held by the director and not shared by the stockholders;[58] and the duty of care, which requires directors to make reasonable and informed decisions.[59] Section 102(b)(7) of Delaware's General Corporation Law permits charter provisions that protect directors from monetary liability for breach of the duty of care, but not for breach of the duty of loyalty or for intentional misconduct or bad faith.[60] One might imagine that this lack of protection for disloyalty and the like might open the door to holding directors liable for the results of artificial intelligence used by the corporation. However, the standard for proving disloyalty, intentional misconduct, or bad faith is high.[61] Although "an utter failure" to

54. Tellabs, Inc. v. Makor Issues & Rts., Ltd., 551 U.S. 308, 314 (2007).

55. *See id.* at 323 ("[I]n determining whether the pleaded facts give rise to a "strong" inference of scienter, the court must take into account plausible opposing inferences. The Seventh Circuit expressly declined to engage in such a comparative inquiry.").

56. *Corporations*, CORNELL L. SCH. LEGAL INFO. INST., https://perma.cc/5EQW-K9GQ (archived Nov. 7, 2021).

57. Of all states, Delaware's corporate laws are perhaps the most relevant to examine. They are generally regarded as the gold standard among corporation laws, and benefit from a Chancery Court that specializes in corporate law disputes. As a result, many large companies choose to incorporate in Delaware. *See* Jan Ting, *Why Do So Many Corporations Choose to Incorporate in Delaware?*, WHYY (Apr. 27, 2011), https://perma.cc/BD7E-9DYN.

58. *See In re* Oracle Corp. Derivative Litig., No. 2017-0337-SG, 2018 WL 1381331, at *11 (Del. Ch. Mar. 19, 2018) ("The duty of loyalty requires directors to put the best interests of the corporation ahead of any other interest held by the directors and not shared by the stockholders.").

59. Brehm v. Eisner, 746 A.2d 244, 258-59 (Del. 2000) ("[I]n making business decisions, directors must consider all material information reasonably available, and . . . the directors' process is actionable only if grossly negligent.").

60. DEL. CODE ANN. tit. 8, § 102(b)(7) (West 2020). The exact language of the statute states that corporations may include provisions in their founding charter that limit the liability of directors for breaches of fiduciary duty *except* in cases of breach of the duty of loyalty, intentional misconduct, and bad faith. *Id.*; *see also* Geoffrey P. Miller, *A Modest Proposal for Fixing Delaware's Broken Duty of Care*, 2010 COLUM. BUS. L. REV. 319, 322 (2010) ("Section 102(b)(7) of the General Corporation Law allows Delaware firms to enact charter amendments shielding directors from monetary liability for nonintentional breaches of the duty of care.").

61. *See, e.g.*, *In re* Merge Healthcare, Inc., No. 11388-VCG, 2017 WL 395981, at *1 (Del. Ch. Jan. 30, 2017) (noting that "demonstrati[ng] that the directors breached the duty of loyalty" is "a rather difficult target for a plaintiff to hit," while showing gross negligence "is less formidable than showing disloyalty"); Dawson v. Pittco Cap. Partners, L.P., No. 3148-VCN, 2012 WL 1564805, at *33 (Del. Ch. Apr. 30, 2012) ("[I]ntentional misconduct remains

exercise oversight constitutes a breach of the duty of loyalty,[62] such "*Caremark*" claims are notoriously hard to prove.[63] Moreover, directors are protected by 102(b)(7) charter provisions from monetary liability for gross negligence and even recklessness, both of which violate the duty of care but not the duty of loyalty.[64] Nor is gross negligence or recklessness sufficient to establish bad faith.[65] In short, with respect to the results of artificial intelligence used by corporations, directors are unlikely to be proved liable for disloyalty, intentional misconduct, or bad faith, and are protected by 102(b)(7) charter provisions from monetary liability for gross negligence and recklessness.

In addition to holding officers and directors of a corporation liable for securities fraud, the corporation itself may be liable under federal and state securities law. Corporate wrongdoing is a function of actions taken by its human employees, and courts differ regarding the way in which liability may be imputed from employees to a corporation. As a general matter, courts rely on the principle of respondeat superior, which holds an employer liable for the actions of its employees.[66] This approach draws from the state of mind of individual decision-makers within the company to impute corporate scienter.[67] Certain circuits, however, have allowed plaintiffs to infer corporate

---

a difficult standard to meet . . . . *[I]ntentional misconduct* must be proven. This requires a showing that the fiduciary intentionally harmed those to whom the duty is owed."); McElrath v. Kalanick, 224 A.3d 982, 992 n.46 (Del. 2020) (noting "the difficulties in proving bad faith director action" (quoting City of Birmingham Ret. & Relief Sys. v. Good, 177 A.3d 47, 55 (Del. 2017))).

62. *In re* Caremark Int'l Inc. Derivative Litig., 698 A.2d 959, 971 (Del. Ch. 1996) ("[A] sustained or systematic failure of the board to exercise oversight—such as an utter failure to attempt to assure a reasonable information and reporting system exists—will establish the lack of good faith that is a necessary condition to liability.").

63. *McElrath*, 224 A.3d at 992 n.46 (noting that "a *Caremark* claim is 'possibly the most difficult theory in corporation law upon which a plaintiff might hope to win a judgment'" (quoting City of Birmingham Ret. & Relief Sys. v. Good, 177 A.3d 47, 55 (Del. 2017))).

64. *In re* Walt Disney Co. Derivative Litig., 906 A.2d 27, 64-65 (Del. 2006) ("[W]e address the issue of whether gross negligence (including a failure to inform one's self of available material facts), without more, can also constitute bad faith. The answer is clearly no."); McPadden v. Sidhu, 964 A.2d 1262, 1275 (Del. Ch. 2008) ("Plaintiff has not, however, sufficiently alleged that the Director Defendants acted in bad faith through a conscious disregard for their duties. Instead, plaintiff has ably pleaded that the Director Defendants quite clearly were not careful enough in the discharge of their duties—that is, they acted with gross negligence or else reckless indifference. Because such conduct breaches the Director Defendants' duty of care, this violation is exculpated by the Section 102(b)(7) provision in the Company's charter and therefore the Director Defendants' motion to dismiss for failure to state a claim must be granted."); *see also Brehm*, 746 A.2d at 259. This contrasts with Rule 10b-5, which does accept recklessness as grounds for liability in some cases. *See* Sundstrand Corp. v. Sun Chem. Corp., 553 F.2d 1033, 1044 (7th Cir. 1977).

65. *Walt Disney*, 906 A.2d at 64-65; *McPadden*, 964 A.2d at 1275.

66. *See* Bradley J. Bondi, *Dangerous Liaisons: Collective Scienter in SEC Enforcement Actions*, 6 N.Y.U. J.L. & BUS. 1, 5-6 (2009) (providing an overview of the respondeat superior approach).

67. *See, e.g.*, Southland Sec. Corp. v. INSpire Ins. Sols., Inc., 365 F.3d 353, 366 (5th Cir. 2004) ("For purposes of determining whether a statement made by the corporation was made by it with the requisite Rule 10(b) scienter we believe it appropriate to look to the state of

scienter collectively, without having to trace the scienter back to an individual employee.[68] Nevertheless, collective scienter is held to an exceptionally stringent standard for those courts that do allow it,[69] and successful pleadings under this standard are uncommon.[70] Consequently, pleading corporate scienter is more reliably alleged by imputing the scienter of relevant, upper-level employees.

Given that demonstrating corporate liability almost always requires some type of individual scienter, existing frameworks make it almost impossible to hold a corporation liable under the securities laws for the damages wrought by its proprietary artificial intelligence software. As explained below in Part II.C., the AI developer may be hard to pinpoint or hold liable because the interconnectivity of different algorithms in the market makes it difficult to determine where one algorithm ends and another begins. Furthermore, the developer may not be at fault within the chain of command given that the AI's self-learning capacity can allow it to act in ways unintended and unforeseen by the programmer. As artificial intelligence becomes increasingly relevant to corporate operations, the law may need to develop more holistic interpretations of scienter, such as the collective approach, in prosecuting artificial intelligence.

Of course, scienter is not always required within the laws related to capital markets. For example, Section 206 of the Investment Advisers Act of 1940 prohibits investment advisors from engaging in "any transaction, practice, or course of business which

---

mind of the individual corporate official or officials who make or issue the statement (or order or approve it or its making or issuance, or who furnish information or language for inclusion therein, or the like) rather than generally to the collective knowledge of all the corporation's officers and employees acquired in the course of their employment.").

68. Bondi, *supra* note 66, at 10-11 ("The strong version posits that corporate scienter is a distinct state of mind that is entirely separate from that of the corporation's individual officers, directors, and employees. Corporate knowledge is an undifferentiated aggregation of its employees' knowledge. Under the strong version of collective scienter, plaintiffs may allege scienter on the part of a corporate defendant without pleading scienter as to any particular employee."); *e.g.*, Teamsters Loc. 445 Freight Div. Pension Fund v. Dynex Cap. Inc., 531 F.3d 190, 196 (2d Cir. 2008) ("Congress has imposed strict requirements on securities fraud pleading, but we do not believe they have imposed the rule urged by defendants, that in no case can corporate scienter be pleaded in the absence of successfully pleading scienter as to an expressly named officer."); *see also* Glazer Cap. Mgmt., LP v. Magistri, 549 F.3d 736, 744 (9th Cir. 2008) ("[T]here could be circumstances in which a company's public statements were so important and so dramatically false that they would create a strong inference that at least *some* corporate officials knew of the falsity upon publication").

69. *E.g.*, Jackson v. Abernathy, 960 F.3d 94, 96 (2d Cir. 2020) (per curiam) ("But while Jackson's allegations support a strong inference that those employees knew of issues with the surgical gown, Jackson has not alleged facts sufficient to impute their knowledge to the corporate entities. And because Jackson has otherwise failed to plead facts tending to show that senior executives must have known that the challenged statements were false, we conclude that Jackson's proposed amended complaint does not raise a strong inference of collective corporate scienter."). In effect, *Jackson* limits which employees may be drawn on as sources to allege collective scienter.

70. *See* Bondi, *supra* note 66, at 11 ("Although no circuit court yet has approved the strong version of collective scienter to establish liability, three circuits seem to have permitted plaintiffs to use the strong version at the pleading stage to allege scienter.").

operates as a fraud or deceit."[71] Recent decisions demonstrate that, although some clauses of Section 206 require showing scienter, mere negligence is sufficient to show wrongdoing under Section 206(2).[72] Although limited to the context of investment advisors, Section 206(2) could provide an approach for addressing improper activity involving artificial intelligence. For example, if an artificial intelligence managing a client's investment portfolio suddenly converted the entire pool of assets to Dogecoin, which then plunged in value, the client likely cannot show that the investment adviser using the artificial intelligence *intended* to deceive the client about investing her money in a diversified portfolio. However, the client may be able to show that the adviser acted negligently by not having safeguards or audits to prevent the algorithm from straying from the promised investment strategy.

Other provisions hold actors to a strict liability standard. For example, Section 11 of the Securities Act of 1933, which regulates securities registration statements such as the information in a company's initial public offering (IPO), applies a strict liability standard.[73] Strict liability does not require demonstrating intent or scienter,[74] which makes it easier to prove wrongdoing under Section 11 compared to, say, Rule 10b-5 or corporate governance laws. Rather, negligent errors in a registration statement—for example, those caused by artificial intelligence—could be grounds for Section 11 liability.[75] As with Section 206 of the Advisers Act, however, Section 11 covers only a narrow part of the landscape—and a shrinking one at that, given that capital is raised increasingly in private markets which are not governed by Section 11.[76]

---

71. 15 U.S.C. § 80b-6.

72. *See* Robare Grp., Ltd. v. SEC, 922 F.3d 468, 472 (D.C. Cir. 2019) ("A violation of Section 206(1) requires proof of 'scienter,' that is, proof of an 'intent to deceive, manipulate, or defraud.' Proof of simple negligence suffices for a violation of Section 206(2), however." (citation omitted)). *Compare* 15 U.S.C. § 80b-6(1) ("employ any device, scheme, or artifice to defraud any client or prospective client"), *with* 15 U.S.C. § 80b-6(2) ("engage in any transaction, practice, or course of business which operates as a fraud or deceit upon any client or prospective client").

73. 15 U.S.C. § 77k(a).

74. *See* SECURITIES ACT OF 1933, LEGAL INFO. INST., https://perma.cc/X4VA-H4X9 (archived Nov. 19, 2021) ("Section 11 makes issuers strictly liable for registration statements that contain 'an untrue statement of a material fact or omit to state a material fact required . . . to make the statements there in no misleading.' . . . As long as the purchaser can trace the purchase back to the initial offering and is within the statute of limitations, he can sue—there is no need to prove causation or reliance on the misstatements or omissions.").

75. In contrast, the same negligent errors disclosed in a 10-K would not be subject to liability under Rule 10b. *See supra* text accompanying notes 48-50. One scholar has argued that, in obviating the need to demonstrate scienter, strict liability may be better suited for governing emerging and risky technologies like AI because it provides a stronger incentive toward the further development of safety measures. Herbert Zech, *Liability for AI: Public Policy Considerations*, 22 ERA FORUM 147, 152-153 (2021).

76. *See* Elisabeth de Fontaney, *The Deregulation of Private Capital and the Decline of the Public Company*, 68 HASTINGS L.J. 445, 453-455, 467 (2017) (describing the way in which declines in the number of IPOs and listings on public exchanges have been matched by increased capital-raising in private, unregulated markets); *see also* Elizabeth Pollman, *Private Company Lies*, 109 GEO L.J. 353, 353 (2020) (noting that the legal and regulatory regimes have

Generally, existing measures that determine liability and wrongdoing in the financial realm depend on demonstrating intent and motivation—qualities conspicuously absent from modern forms of artificial intelligence. For example, the "intent" may lie in decisions made by the artificial intelligence itself, long after the creators or even the users have let go of the reins. The disconnect between current legal standards and the character of artificial intelligence offers a troubling amount of room for algorithms to operate without accountability.

II. Shades of Things to Come

As the centrality of intent to financial regulation demonstrates, existing legal regimes are not structured to deal with the problems introduced by the growing use of artificial intelligence. Recent episodes in the financial industry that involve adjacent technologies augur the potential for serious artificial intelligence-induced disruptions down the line.

A. Who's Responsible?

When an artificial intelligence develops its own strategies for problem-solving and acts in ways unforeseen by its maker, who is responsible for its actions? Several real and hypothetical examples suggest how these situations may engender conflict with current regulatory systems.

Consider, first, the *Coscia* case, in which an unscrupulous trader was convicted of "spoofing" in the copper commodities market.[77] Spoofing is a market manipulation technique in which traders use algorithms to submit futures orders, then cancel or modify the orders before they are executed.[78]

Imagine an unscrupulous trader who wishes to manipulate the market in cattle futures. The spoofer places a series of large orders, which will be cancelled before execution, along with a smaller order at the price at which the spoofer really wants to buy or sell. Using extremely rapid trading strategies[79] that require utilization of a

been oriented mainly toward public markets, at risk to investor safety).

77. *See* United States v. Coscia, 866 F.3d 782 (7th Cir. 2017).

78. *Spoofing*, CORP. FIN. INST., https://perma.cc/5AFH-DTME (archived Nov. 19, 2021); *see also* Gina-Gail S. Fletcher, *Deterring Algorithmic Manipulation*, 74 VAND. L. REV 101, 134-135 (2021) (explaining the process of spoofing and how high-frequency trading is well-suited to magnify the practice).

79. To be sure, the practice of high-frequency trading (HFT) is completely legal when the executed trades themselves are legitimate. Some proponents of the practice even claim that legitimate HFT benefits financial markets by increasing liquidity and efficiency. *See generally* Charles P. Henness & Ben Niu, *An Overview of High-Frequency Trading*, RMB CAP. (Jun. 16, 2020), https://perma.cc/25AC-CCKJ. Other commentators, however, have highlighted as problems the lack of transparency—both in high-frequency algorithms' operation and the appearance of investment "dark pools"—in addition to the institutional advantage that HFT confers to large players in the market. *See, e.g.*, Michael J. McGowan, *The Rise of Computerized High Frequency Trading: Use and Controversy*, 9 DUKE L. & TECH. REV. 1, 20-22 (2010).

trading algorithm because they could not be accomplished physically by a human, the spoofer can make it appear that the market is receiving orders at a continually decreasing price. For example, if a series of large orders comes in at decreasing values of $3.05, then $3.04, then $3.03, it will appear that the market is littered with individuals willing to sell at lower and lower prices.[80] When the supposed price hits $3, the spoofer's buy order of $3 will be executed. All of this will happen in an eye blink, during which time the decreasing value orders will be cancelled. The spoofer, now holding futures at a cost of $3, can employ the strategy in reverse to drive the price back up and sell at a profit.[81]

Testimony in the *Coscia* case established that the spoofer directed a programmer to develop an algorithm, intended to operate as a decoy to pump the market and induce other trading algorithms within the market to operate in a certain manner.[82] Suppose, however, that the trading program was simply an artificial intelligence directed to develop and execute a strategy that would optimize investment returns in the market for cattle futures. The machine learning program itself developed the spoofing as a method of reaching that goal. In theory, an ethical programmer could implement safeguards. In reality, our experience with artificial intelligence systems suggest that they can develop in ways that are unintended and unpredictable.

Prosecutors would find it difficult to obtain a conviction. Legal standards in the realm of market manipulation generally require proof that the defendant intended to manipulate the market.[83] Unless one contemplated putting the artificial intelligence on trial, the "defendant" would be the unscrupulous trader, for whom one cannot prove any specific intent. What could be wrong in asking the artificial intelligence to make a

80. *See Coscia*, 866 F.3d at 787.

81. *See id.*

82. *See id.* at 789.

83. *See, e.g.*, Braman v. The CME Grp., Inc., 149 F. Supp. 3d 874, 890 (N.D. Ill. 2015) ("Plaintiffs have not alleged that the defendants intended to cause artificial prices, or even that they intended to cause the HFTs to behave in ways that would artificially affect prices. . . . Because high frequency trading itself does not violate the CEA, these allegations are not enough to establish any specific intent on the part of the defendants."); SEC v. Masri, 523 F. Supp. 2d 361, 372 (S.D.N.Y. 2007) ("The Court concludes, therefore, that if an investor conducts an open-market transaction with the intent of artificially affecting the price of the security, and not for any legitimate economic reason, it can constitute market manipulation. Indeed, 'the only definition [of market manipulation] that makes any sense is subjective—it focuses entirely on the intent of the trader.'" (quoting Daniel R. Fischel & David J. Ross, *Should the Law Prohibit "Manipulation" in Financial Markets?*, 105 HARV. L. REV. 503, 510 (1991))); Santa Fe Indus. Inc. v. Green, 430 U.S. 462, 473-74 (1977) ("'When a statute speaks so specifically in terms of manipulation and deception, . . . and when its history reflects no more expansive intent, we are quite unwilling to extend the scope of the statute . . . .' Thus the claim of fraud and fiduciary breach in this complaint states a cause of action under any part of Rule 10b-5 only if the conduct alleged can be fairly viewed as 'manipulative or deceptive' within the meaning of the statute." (citation omitted) (quoting Ernst & Ernst v. Hochfelder, 425 U.S. 185, 212-214 (1976))); *see also* Yavar Bathaee, *The Artificial Intelligence Black Box and the Failure of Intent and Causation*, 31 HARV. J.L. & TECH. 890 (2018) (discussing the difficulty of developing evidence of inappropriate use of high-frequency-trading algorithms).

profit?

To be sure, spoofing is prosecuted differently depending on the market being manipulated. Government agencies have traditionally pursued spoofing cases in the securities market under Section 9(a)(2) of the 1934 Exchange Act and Sections 10(b) and 17(a) of the 1933 Securities Act.[84] Spoofing has not yet been mentioned by name in securities legislation; it is instead usually classed as a type of "open-market manipulation"[85] and accordingly requires, at the very least, the intent to artificially raise or depress the market price.[86] The practice was criminalized for commodities markets under different terms in the 2010 Dodd-Frank Wall Street Reform and Consumer Protection Act in Section 6c(a)(5).[87] This section, which is specific to spoofing, no longer requires proving the intent to create an artificial price. Instead, the Commodity Futures Trading Commission (CFTC) or Department of Justice (DOJ) merely needs to prove the claimant intended to cancel her entire bid or offer.[88] Although this distinction makes it significantly easier to successfully bring charges against a human spoofer, it does not change the difficulty of charging an artificial intelligence. Even if an artificial intelligence were to "intentionally" cancel a large number of bids in its efforts to create more favorable trading conditions for itself, the artificial intelligence itself could not be charged with intentional wrongdoing. So long as the artificial intelligence was not designed to spoof, moreover, neither could the courts charge its maker or user under this statute.

Spoofing is today's bogeyman, but what about the techniques that have yet to be developed? A trader could carefully program the artificial intelligence to scrupulously observe all legal systems, including avoiding any spoofing. If the artificial intelligence develops a method of manipulating market pricing in a manner that no one could have contemplated, the trader certainly would have no specific intent. Thus, the problem could occur even with a completely honest trader.

Of course, the desire to maximize profits in stock trading can lead to troubling results if unbounded. Consider a potential variant on the recent Robinhood case. Robinhood Financial LLC markets itself as a broker-dealer for novice investors. The company launched its mobile trading platform in 2015, attracting clients by waiving

84. *See* Jessica Masella & Jonathan Haray, *'Spoofing' Prosecutions: The DOJ's Approach*, N.Y. L. J. (Apr. 2, 2021), https://perma.cc/UT8Z-QNW5.

85. *Id.*

86. *See, e.g.*, Markowski v. SEC, 274 F.3d 525 (D.C. Cir. 2001) (holding that open-market trades that are otherwise legal can be considered illegal manipulation if done with the intent to raise or depress the market price); *cf.* ATSI Commc'ns v. Shaar Fund, Ltd., 493 F.3d 87, 100-01 (2d Cir. 2007) (holding that a trade cannot be considered illegal manipulation unless the transaction is "willfully combined with something more to create a false impression of how market participants value a security").

87. 7 U.S.C. § 6c(a)(5); *see also* Masella & Haray, *supra* note 84.

88. *See* Masella & Haray, *supra* note 84; United States v. Coscia, 177 F. Supp. 3d 1087, 1094 (N.D. Ill. 2016) (holding that the government must prove that "at the time Mr. Coscia entered the bid or offer . . . he intended to cancel the entire bid or offer before it was executed").

the commission that most broker-dealers at the time charged per transaction.[89] With zero commissions, Robinhood made money instead by sending its orders to broker-dealers such as Citadel LLC in return for payments from these firms.[90] As a result, Robinhood customers ultimately received execution prices inferior to those provided by other brokers.[91] Robinhood not only failed to disclose its reliance on "payment for order flow" for revenue; it also claimed that "its execution quality matched or beat that of its competitors," when in fact its prices "deprived customers of $34.1 million *even after* taking into account the savings from not paying a commission."[92]

In December 2020, the SEC charged Robinhood for failure to disclose its use of payment for order flow and for failure to satisfy its duty of best execution.[93] The company settled the case for $65 million.[94] Given Robinhood's stated purpose to make finance accessible to all, in failing to disclose to its customers the company's actual source of revenue, Robinhood effectively aligned itself with the Wall Street firms it proclaimed to reject.

As with the spoofing hypothetical above, suppose a broker-dealer had not intentionally designed its systems to operate in a manner that obtained revenue by providing poorer executions for its customers. Suppose, instead, that the company had merely instructed an artificial intelligence to design the most efficient and profitable way of executing large numbers of small trades for customers. If the artificial intelligence had developed this, or even more complex and nefarious systems, where would scienter lie? Who would be to blame? And can or should this responsibility be able to be "disclosed away"?

These types of concerns are more than theoretical, given that artificial intelligence systems have demonstrated a remarkable ability to cheat in ways that are entirely unpredictable to those who created the program. For example, a team of Google and Stanford researchers attempted to train an artificial intelligence to transform aerial photographs into street maps as well as from street maps back to aerial photographs.[95] The researchers found that, instead of learning how to reconstruct the aerial photographs, the network learned how to cheat by hiding information that it would later need in a "nearly imperceptible high-frequency signal."[96] As the title of one press article

89. Sheelah Kolhatkar, *Robinhood's Big Gamble*, NEW YORKER (May 17, 2021), https://perma.cc/Z8U7-D29E.

90. Press Release, SEC, SEC Charges Robinhood Financial With Misleading Customers About Revenue Sources and Failing to Satisfy Duty of Best Execution (Dec. 17, 2020), https://perma.cc/7W3U-9XRC.

91. *Id.*

92. *Id.* (emphasis added).

93. *Id.*

94. *Id.*

95. *See* Devin Coldewey, *This Clever AI Hid Data from Its Creators to Cheat at Its Appointed Task*, TECHCRUNCH (Dec. 31, 2018), https://perma.cc/UEH8-EY36; *see also* Casey Chu, Andrey Zhmoginov, & Mark Sandler, *CycleGAN, a Master of Steganography* (31st Conference on Neural Information Processing Systems, 2017), https://arxiv.org/pdf/1712.02950.pdf.

96. *See* Coldewey, *supra* note 95.

explained, the artificial intelligence figured out how to cheat at the programmed activity by hiding data from its creators that could later be used to enhance its ability to perform the task.[97] As with the Berlin smart house that engaged in a denial of service attack,[98] this example highlights the potential for artificial intelligence systems to operate in ways that humans would consider cheating or inappropriate behavior to get to the desired result.

Although one might imagine that an artificial intelligence system could be programmed to avoid the problem, it is extremely difficult for programmers to predict—and sometimes even understand—the eventual choices made by the artificial intelligence as it operates. For example, in a project conducted by the research company OpenAI, artificial intelligence agents learned to manipulate objects in a hide-and-seek game in ways the researchers did not realize were physically possible.[99] As one member of the research team noted, "[t]he agents will come up with these unexpected behaviors, which will be a safety problem down the road when you put them in more complex environments."[100]

Consider another example. Imagine a semi-autonomous car with features such as automatic braking when a pedestrian is detected on the road. Suppose the car's artificial intelligence engages automatic braking when there is no pedestrian. The braking causes a skid, and the driver is killed—perhaps even along with passengers in other cars. The action has not occurred in any simulations and does not follow a logical path that the programmers understand. Are the company and its executives off the hook for any liability under the securities law?[101]

One could argue that the semi-autonomous car example is no different than when any other vehicle system malfunctions in an unanticipated manner, ranging from the accelerator to the airbag. The question, however, is what constitutes "unanticipated" in the context of artificial intelligence. To some extent, the point of artificial intelligence is to develop new approaches in a way that is more effective than human intelligence can manage. Yet that can lead to circumstances beyond what human intelligence can *anticipate*. Who then is responsible? Does a company simply say, "We use AI. Don't blame us for anything that happens"? Alternatively, must they simply stop using artificial intelligence? Or is there a better approach?

In certain circumstances, such as an IPO filing,[102] the law holds the company to a

97. *See id.*

98. *See* Hill, *supra* note 32.

99. *See* Eliza Strickland, *AI Agents Startle Researchers with Unexpected Hide-and-Seek Strategies*, IEEE SPECTRUM (Sep. 17, 2019), https://perma.cc/SP8K-LZQZ.

100. *See id.*; *see also* Warren & Hillas, *supra* note 35 (describing the difficulty of explaining an AI's choices in the context of debates surrounding lethal autonomous weapons systems); Hill, *supra* note 32 (describing a smart house engaging in a denial-of-service attack).

101. We note that liability under tort or other legal provisions would be a separate question beyond the scope of this article.

102. 15 U.S.C. § 77k(a); *see* SECURITIES ACT OF 1933, *supra* note 74 ("Section 11 makes issuers strictly liable for registration statements that contain 'an untrue statement of a material fact or omit to state a material fact required . . . to make the statements there in no

standard of strict liability. In circumstances without strict liability, however, where scienter must be proven, proof will be difficult to develop when the harm in question arises unforeseen and unintended. Consequently, finding the company liable creates challenges when the source of wrongdoing is a program and not its programmer or operator.

This variety of scenario is not strictly hypothetical. As described above,[103] even machines make mistakes. The Knight Capital software glitch hints at the magnitude of losses that can take place through simple errors.[104] On the morning of August 1, 2012, an errant and untested section of code added to Knight Capital's automated trading router caused the firm to issue more than 400 million transaction orders (instead of the 212 customer orders it meant to handle).[105] In the forty-five minutes of mistaken trading, Knight—one of the premier high-speed trading firms at the time[106]—amassed several billion dollars' worth of unwanted positions, which translated to about $460 million in losses, more than $10 million per minute.[107] The effects of the error reverberated far beyond the company itself. The algorithms generated volatile price swings in nearly 150 stocks that morning.[108]

Similarly, scholars have documented rapid spikes and crashes occurring across the broad market and within individual stocks, including more than 18,000 over a five-year period, each lasting about 1,500 milliseconds.[109] In slightly longer mini-crash episodes, which have plagued companies including Apple and Citicorp, billions of dollars of capitalization have been wiped out and recovered in a few minutes.[110] The stocks

misleading.' . . . As long as the purchaser can trace the purchase back to the initial offering and is within the statute of limitations, he can sue—there is no need to prove causation or reliance on the misstatements or omissions.").

103. *See, e.g.*, *supra* text accompanying notes 30-36 (examples of contemporary artificial intelligence errors).

104. *See generally* Yadav, *supra* note 11, at 1047 (describing the Knight Capital mishap).

105. Press Release, SEC, SEC Charges Knight Capital with Violations of Market Access Rule (Oct. 16, 2013), https://perma.cc/SEW8-KUR3 [hereinafter SEC Knight Capital Press Release].

106. *See* Nathaniel Popper, *Knight Capital Says Trading Glitch Cost It $440 Million*, N.Y. TIMES (Aug. 2, 2012, 9:07 AM), https://perma.cc/QD5G-BYUW (noting that Knight Capital enacted 11% of all American stock trades between January and May 2012).

107. SEC Knight Capital Press Release, *supra* note 105. To compound its colossal losses that morning, Knight Capital had the distinction of receiving the first citation of the SEC's market access Rule 15c3-5, which requires broker-dealers to provide safeguards that mitigate the risks created by their access to the market.

108. *See* Melanie Rodier, *Reliving the Knightmare: Where Were the Risk Controls?*, 30 WALL ST. & TECH. (2012) (describing how an error in Knight Capital's untested trading algorithm, in addition to costing the firm $460M, also generated price swings of as much as 10% in 148 stocks on the NYSE).

109. Neil Johnson, Guannan Zhao, Eric Hunsader, Hong Qi, Nicholas Johnson, Jing Meng & Brian Tivnan, *Abrupt Rise of New Machine Ecology Beyond Human Response Time*, 3 NAT. SCI. REPS. 2627 (Sept. 11, 2013), https://perma.cc/PAD4-VA5M, *reprinted in* Yadav, *supra* note 11, at 1077.

110. *See* Phillip Elmer-DeWitt, *Snapshot of an Apple Flash Crash*, FORTUNE (Feb. 11, 2011, 4:50 PM), https://perma.cc/4BMW-992P; Graham Bowley, *The Flash Crash, in Miniature*,

may recover, but individual traders and trades can be harmed in the maelstrom.

Artificial intelligence systems are designed to learn by testing and iteration. Suppose in its process of testing and iteration, an artificial intelligence system creates an error, triggering a cascade similar to the Knight Capital incident. The consequences to the company using the artificial intelligence, not to mention to individuals and financial markets themselves, could be dramatic. Once again, who would we blame?

Moreover, when does the simple existence of errors become a predictable outcome in itself? And if we reach that point, on the whole or for a specific category of error, who should we hold accountable?

B. Size Doesn't Matter

Regulation of actors in the securities markets often revolves around size. For example, the SEC's Regulation Systems Compliance and Integrity (Reg SCI) governs the electronic and IT systems of a few dozen large entities.[111] These include stock exchanges and trading platforms, on which financial markets operate.[112] Reg SCI became effective in 2015 and outlines expectations regarding the capacity, integrity, resiliency, availability, and security in order that electronic and IT systems can facilitate fair markets.[113] Specific requirements include stress testing and the adoption of standards to ensure reliable operability.[114] Reg SCI also requires these large entities to report any system disruptions to the SEC and undertake any corrective action to mitigate investor harm.[115] In March 2018, the agency issued its first violations under the order, charging the New York Stock Exchange for wrongfully imposing a trading halt and applying price collars without having requisite rules in place.[116]

The SEC intended the regulation to help improve market resilience and the oversight capacity of the agency,[117] designs that may be compromised by the rule's limited scope. Although Reg SCI covers entities that are core to the market, infrastructural stability and operability may also be compromised by the smaller entities allowed to operate outside of the rule's purview. These excluded parties include smaller trading venues and thousands of broker-dealers, in addition to most of the electronic and IT

N.Y. TIMES (Nov. 8, 2010), https://perma.cc/4747-AEWZ , *reprinted in* Yadav, *supra* note 11, at 1078.

111. When the regulation was enacted in 2014 (and went into effect in 2015), it covered 44 entities. *See* Commissioner Kara M. Stein, Public Statement on Regulation Systems Compliance & Integrity (SCI) (Nov. 19, 2014), https://perma.cc/K3N3-TK6X [hereinafter Commissioner Stein Public Statement].

112. Regulation Systems Compliance & Integrity, 17 C.F.R. § 242.1001.

113. *Id.*

114. *Id.*

115. Regulation SCI, 17 C.F.R. § 242.1002.

116. Press Release, SEC, NYSE to Pay $14 Million Penalty for Multiple Violations (Mar. 6, 2018), https://perma.cc/FF3C-LMFK.

117. Chair Mary Jo White, Statement at Open Meeting on Regulation SCI (Nov. 19, 2014), https://perma.cc/R4Z5-D886.

systems of most market players.[118] Beyond broker-dealers, Reg SCI also stops short of scrutinizing the activity of smaller, intraday proprietary trading firms that rely on algorithms.[119] Given their high transaction volume and the interconnected, data-driven quality of financial markets, rogue behavior from these market participants can potentially also generate dangerous waves across the whole market.[120]

The law also assumes that size matters for raising capital in non-public markets. SEC regulations allow what are usually smaller companies to raise funds from investors outside of public markets. Permitting companies to raise capital without registering with the SEC is allowed by exemptions authorized by Congress or the SEC. Many of these private offering exemptions are premised on the principle that the offering will be made only to wealthy or sophisticated investors. Specifically, Regulation D's safe harbor for raising money through a Rule 506 private offering[121] requires that such an offering be made almost exclusively to "accredited investors," a designation granted to individuals meeting a certain income or net worth threshold.[122] Under the rule, an accredited investor is a person with more than $200,000 in yearly earnings for at least two years, or a net worth exceeding $1 million.[123] This private offering exemption also allows for the offering to be made to up to thirty-five non-accredited investors who meet certain sophistication requirements such as financial expertise—CPAs or bank executives, for example.[124] But by and large, private market regulation replicates the focus of public market regulation, allowing exemptions for large, wealthy, experienced, and sophisticated players.

In the modern context of artificial intelligence, however, size does not matter. One must think of artificial intelligence as a capacity that can enhance the abilities of those who use it. Just as the power of flight opens new avenues for superheroes, so does the power of artificial intelligence open new horizons for mere mortals. One can already see that potential in play. The "flash crash" of May 6, 2010, for example, provides a

118. *See* Commissioner Stein Public Statement, *supra* note 111.

119. *Id.*

120. *See infra* text accompanying notes 147-150.

121. *See* Regulation D, 17 C.F.R. § 230.506 (2021). *See generally Private Placements-Rule 506(b)*, SEC, https://perma.cc/D4PE-HVRH (archived Nov. 22, 2021). Rule 506(b) of Regulation D allows companies to meet the requirements of Section 4(a)(2) of the Securities Act. This safe harbor provision allows private companies to raise an unlimited amount of money from accredited investors (and up to thirty-five non-accredited, "sophisticated" investors). See notes 122-124 below for definitions of accredited and sophisticated investors.

122. *See* 17 C.F.R. § 230.501(a) (defining accredited investors under Rule 506(b)).

123. *Id.* Accredited investor status can also be gained by meeting other thresholds, such as, for licensed corporations and trusts, an asset total exceeding $5 million. *Id.*

124. *See* 17 C.F.R. § 230.506(b)(2)(ii) (defining sophisticated investor as "[e]ach purchaser who is not an accredited investor either alone or with his purchaser representative(s) has such knowledge and experience in financial and business matters that he is capable of evaluating the merits and risks of the prospective investment"); *see also Rule 506 of Regulation D*, INVESTOR.GOV, https://perma.cc/CP4L-KAJP (archived Nov. 22, 2011) (noting that, along with accredited investors, private companies may raise money from up to thirty-five sophisticated investors, so long as they are furnished with additional disclosure information).

warning of the potential influence wielded by rogue individuals equipped with trading software. On that day, the Dow Jones index fell 5-6% in a span of a few minutes, with many securities trading at prices 60% or more away from their value just moments earlier.[125]

Although the government's investigation pointed to one large, automated trade as the trigger of the flash crash,[126] authorities also charged a British high-frequency trader, operating alone, with issuing a surge of "spoof orders"[127] that may have contributed to the extreme market volatility.[128] Navinder Singh Sarao, the convicted trader, customized an automated program that exploited the lockstep movement of other high-frequency traders in the market, a product of their similarly programmed trading software.[129] In simple terms, Sarao's "spoof orders" prompted a surge of trading in one direction that Sarao then profited off of through genuine buy or sell orders.[130] Sarao's spoof orders exacerbated volatile conditions in the market, increasing the magnitude of swings in supply and demand.[131] Thus, relatively homogeneous automated trading patterns contributed to a brittle marketplace, rendering it vulnerable to the manipulations of a single individual. Artificial intelligence enhances the dangers, provides an additional layer of distance, and allows for new types of defenses based on lack of intent or control.

To be sure, the orders that helped precipitate the "flash crash" were deliberately programmed, not the product of autonomous or self-improving artificial intelligence software.[132] The specter of more advanced, more opaque and more complicated artificial intelligence only heightens the potential danger of rogue individuals in the

125. U.S. COMMODITY FUTURES TRADING COMM'N & SEC, FINDINGS REGARDING THE MARKET EVENTS OF MAY 6, 2010 1 (2010), https://perma.cc/E7CF-X8L5 [hereinafter CFTC & SEC Report].

126. *Id.* at 2-3.

127. Spoofing is a practice in which traders use algorithms to submit futures buy/sell orders, then canceling or modifying the orders before they are executed in order to manipulate market prices. *See supra* text accompanying notes 77-81; *Spoofing*, *supra* note 78; *see also* Fletcher, *supra* note 78, at 134-35 (explaining the process of spoofing and how high-frequency trading is well-suited to magnify the practice).

128. CFTC v. Nav Sarao Futures Limited, No. 15-cv-3398 at 16-18 (N.D. Ill. Nov. 14, 2016) (consent order granting permanent injunction). *But see* Eric M. Aldrich, Joseph A. Grundfest & Gregory Laughlin, *The Flash Crash: A New Deconstruction* (Mar. 26, 2017), https://ssrn.com/abstract=2721922 (finding, in an empirical analysis, that Navinder Singh Sarao's trading behavior minimally impacted market prices on May 6, 2010). The joint CFTC/SEC report did not name Singh, but did suggest that the frenzied activity of high-frequency traders threatened market liquidity, contributing to the crash. *See* CFTC & SEC Report, *supra* note 125, at 14-16.

129. Andy Verity & Eleanor Lawrie, *Hound of Hounslow: Who Is Navinder Sarao, the 'Flash Crash Trader'?*, BBC NEWS (Jan. 28, 2020), https://perma.cc/L6SR-67NF.

130. *Id.*; *see also supra* text accompanying notes 77-81 (outlining how a spoofing scheme functions).

131. *See* CFTC & SEC Report, *supra* note 125, at 14-16 (discussing the impact of high-speed trading in creating conditions conducive to the crash).

132. *See* Fletcher, *supra* note 78, at 112.

marketplace. Although no more difficult to envision, a flash crash driven by deep learning alone would certainly present a greater challenge to diagnose, much less prosecute.

Even more troubling is the so-called "Hash Crash," which occurred in 2013.[133] As part of a hack of the Associated Press' Twitter feed, a false tweet announced an explosion at the White House and suggested that President Obama had been injured or killed.[134] The episode lasted only minutes until the truth emerged and the market rebounded to its original position, but during those minutes, the Dow Jones Industrial Average dropped 145 points, losing 1% of its value.[135]

Suppose that the actor in a tweet-gate of this kind had been an artificial intelligence? To offer a variant of the scenario one scholar has posited, imagine that an artificial intelligence was instructed to study and capitalize on social media trends to maximize profits, then engaged in extensive retweeting of false information—the type of information that frequently circulates on the Internet.[136] In that case, a single player could destabilize markets, and profit in the process.

The emphasis on certain big players such as stock exchanges, albeit limited, is not necessarily wrong. The concept generally falls within the notion described below of touchpoint regulation.[137] The law can operate when the related activity and those who engage in it intersect with the financial system. The law, however, must reach all who connect with the touchpoint, not just the big folks.

Following in the footsteps of the SEC's Reg SCI, the CFTC contemplated, but failed to finally approve, a policy related to trading algorithms. In 2015, the agency proposed Regulation Automated Trading (AT), which would require market participants to register their trading algorithms with the CFTC, sufficiently test algorithms prior to market launch, and implement checks such as kill switches and order cancellation systems.[138] Thus, Regulation AT would, in addition to market exchanges, directly address the actors trading with algorithms, an important constituency of algorithm

133. *See* Chiponda Chimbelu, *Twitter's Hash Crash: Social Media and the Financial Markets*, DW (May 27, 2013), https://perma.cc/8ASR-DHTA.

134. *See id.*; Paloma Migone, '*Hash Crash' Further Erodes Confidence in US Market - Report*, TRADE NEWS (May 7, 2013), https://perma.cc/KG5Z-9PDG.

135. Migone, *supra* note 134; *cf.* Jena McGregor, *Elon Musk's April Fools' Tweets Were 'Not a Joking Matter,' Experts Say*, WASH. POST (Apr. 3, 2018), https://perma.cc/S6T2-KP4V (reporting that in 2018, Tesla shares fell 5% after founder Elon Musk tweeted an April Fools' Day bankruptcy announcement).

136. *See* Bathaee, *supra* note 83, at 911-12 (describing a hypothetical in which a profitable trading algorithm autonomously learns to circulate news articles on Twitter before and after certain trades, helping it move the market and earn a profit). For an example of corrupted social media training data, Microsoft was forced to discontinue its artificial intelligence chatbot after the account, in a matter of hours, began "tweeting" racist and anti-Semitic slurs. Elle Hunt, *Tay, Microsoft's AI Chatbot, Gets a Crash Course in Racism from Twitter*, THE GUARDIAN (Mar. 24, 2016), https://perma.cc/Z6MU-84E7.

137. *See infra* Part III (outlining the concept of touchpoint regulation).

138. Regulation Automated Trading, 80 Fed. Reg. 78,828 (Dec. 17, 2015) (to be codified at 17 C.F.R. pts. 1, 38, 40, 170).

users largely neglected by the SEC's Regulation SCI. However, certain aspects of the proposed rule, such as requiring storage of proprietary algorithms in a repository open to government inspection, raised the hackles of many industry stakeholders.[139]

As a result, the CFTC in 2020 narrowly elected to withdraw the proposed Regulation AT, and instead adopted a more ambiguous set of "risk principles" for electronic trading in the marketplace.[140] CFTC commissioners who voted down Regulation AT in favor of such principles cited the onerous source code disclosure requirement, in addition to the inability of rules to keep pace with rapidly evolving trading technologies.[141]

Ultimately, individual thrusts such as these tackle only parts of the problem. Worse yet, they lack a comprehensive theory for conceptualizing the nature of what we are regulating and the general approach of such regulation. Without a legal framework for what we are doing and why, progress is unlikely, particularly faced with a technology that is rapidly evolving and capable of learning.[142]

C. Where Is It?

Where does an artificial intelligence reside? Where one resides is an exceedingly simple question for most individuals to answer. Even with corporations, the law has developed methods of determining a corporation's home (that is, where it is incorporated) and its presence (where it has sufficient contacts to establish that a court has jurisdiction).[143] With artificial intelligence, however, the question is far murkier—as are

139. *See, e.g.*, Managed Funds Association, Comment Letter on Proposed Regulation Automated Trading (Mar. 16, 2016) at 21, https://perma.cc/MH5X-H8W4 ("Algorithmic managers have all or substantially all of their enterprise value embedded in their code. We are strongly concerned with the proposed requirement that it be made available to regulators and prosecutors upon request, with no need to allege or make a showing of manipulation, fraud or other wrongdoing."). In dissent, one commissioner pointed out that risk principles, in contrast to Regulation AT, may fail to demand meaningful change from market actors. *See* Commissioner Rostin Behnem, Dissenting Statement Regarding Electronic Trading Risk Principles (Jun. 25, 2020), https://perma.cc/U6GY-PV6E. The preamble to the risk principles themselves take note of their redundancy for many stakeholders who already adopted certain risk control measures. *See* Electronic Trading Risk Principles, 85 Fed. Reg. 42,762 (Jan. 11, 2021) (to be codified at 17 C.F.R. pt. 38) ("[T]he Risk Principles may not necessitate the adoption of additional measures by DCMs [i.e. trading markets]."). Consequently, mere risk principles or guidelines may act more as a rubber stamp than an agent of change for the public futures markets.

140. Regulation Automated Trading; Withdrawal, 85 Fed. Reg. 42,755 (Jul. 15, 2020) (to be codified at 17 C.F.R. pts. 1, 38, 40, 170).

141. *See, e.g.*, *id.*; *cf.* Heath P. Tarbert, *Rules for Principles and Principles for Rules: Tools for Crafting Sound Financial Regulation*, 10 HARV. BUS. L. REV. 1 (2020) (discussing the CFTC's history of relying on principle-based rather than rule-based regulation).

142. The need for an effective framework in financial regulation is underscored by the market harms wrought both by balance misinformation and constraining regulatory overreach. *See generally* Yadav, *supra* note 11, at 1033.

143. The notion of personal jurisdiction qualifies a court to decide in a case between two parties. Personal jurisdiction can be satisfied in one of two forms: general jurisdiction and specific jurisdiction. A court has general jurisdiction over a defendant in the location of an individual's residence, or, for corporations, "an equivalent place, one in which the

the spaces inhabited by various forms of artificial intelligence.[144]

If artificial intelligence is an activity, rather than a person or an entity, then one could easily conclude that the artificial intelligence exists wherever the activity is taking place.[145] That answer might be perfectly logical . . . and entirely unhelpful. One could certainly parse through questions related to whether the activity is taking place where the computer user sits (e.g., Ukraine) or where the impact of the behavior lands (the New York Stock Exchange), but it may not help. With artificial intelligence, as with other computer- and internet-related activities,[146] one cannot simply reach into the

---

corporation is fairly regarded as at home." *See* Bristol-Myers Squibb Co. v. Superior Court, 137 S. Ct. 1773, 1780 (2017) (quoting Goodyear Dunlop Tires Operations, S.A. v. Brown, 564 U.S. 915, 924 (2011)). A court with general jurisdiction over an individual or company can take on any claim against that party, even if it occurred outside of the state in question. By contrast, a court only has specific jurisdiction over a nonresident when the suit develops from the defendant's contacts in the forum, or area, the court oversees. *See Bristol-Myers Squibb*, 137 S. Ct. at 1780 (citing Daimler AG v. Bauman, 571 U.S. 117, 127 (2014)); Int'l Shoe Co. v. Washington, 326 U.S. 310, 316 (1945) ("[D]ue process requires only that in order to subject a defendant to a judgment *in personam*, if he be not present within the territory of the forum, he have certain minimum contacts with it such that the maintenance of the suit does not offend 'traditional notions of fair play and substantial justice.'" (quoting Milliken v. Meyer, 311 U.S. 457, 463(1940))).

144. Artificial intelligence threatens to undermine existing tests of personal jurisdiction. In determining whether a court has general jurisdiction over an artificial intelligence, for instance, should the "residence" of the program be the location of the server that stores it? What if the program is stored across several servers for safekeeping? *See* Yvette Joy Liebesman & Julie Cromer Young, *Litigating Against the Artificially Intelligent Infringer*, 14 FIU L. REV 259, 265-266 (2020). Specific jurisdiction is also confused by artificial intelligence, as programs independently operate across many areas, even evolving to act in areas beyond where the artificial intelligence's designer, programmer, or owner intended for it to operate. *See* Zoe Niesel, *Machine Learning and the New Civil Procedure*, 73 SMU L. REV. 493, 534-540 (2020) (outlining how existing tests to determine personal jurisdiction in Internet cases, such as *Zippo Manufacturing Co. v. Zippo Dot Com, Inc.*, 952 F. Supp. 1119 (W.D. Pa. 1997), are apt to fail with an artificially intelligent defendant); *see also* Zoe Niesel, *#personaljurisdiction: A New Age of Internet Contacts*, 94 IND. L.J. 103, 139 (2019) (noting that the attempts of courts to directly translate Internet actions into their physical equivalents are increasingly inappropriate with the advent of artificial intelligence).

145. A cryptocurrency transaction—which occurs across a decentralized distributed ledger network—was decided to be subject to the Exchange Act because the transaction "was validated by a network of global 'nodes' clustered more densely in the United States than in any other country." *In re* Tezos Sec. Litig., No. 17-CV-06779-RS, 2018 WL 4293341, at *8 (N.D. Cal. Aug. 7, 2018).

146. *See generally, e.g.*, Christian A. Camarce, *Harmonization of International Copyright Protection in the Internet Age*, 19 PAC. MCGEORGE GLOBAL BUS. & DEV. L.J. 435 (2007) ("There are numerous challenges posed to the protection of intellectual property rights relating to the nonterritorial nature of the internet."); Lulin Gao, *Intellectual Property Rights in the Internet Era: The New Frontier*, 5 JOHN MARSHALL REV. INTELL. PROP. L. 589 (2006) (describing the history of international intellectual property rights agreements with respect to implications for Internet activity); UNITED STATES COPYRIGHT OFF., INTERNATIONAL COPYRIGHT RELATIONS OF THE UNITED STATES (CIRCULAR 38A) (updated Oct. 2021), https://perma.cc/Z7WA-TETG (listing international intellectual property rights agreements); H. MARSHALL JARRETT, MICHAEL W. BAILIE, ED HAGEN & SCOTT ELTRINGHAM, OFF. LEGAL EDUC., PROSECUTING COMPUTER CRIMES, https://perma.cc/5X8H-UKEJ; Computer

ether to grab ahold of anything or anyone. Nor can one necessarily identify the location of the threat or of those who threaten.

The ability to mask one's activity in the realm of the internet and reach far beyond one's location provides immeasurable challenges, but artificial intelligence compounds these already difficult questions. Algorithms may learn to affect networks indirectly by influencing other actors throughout the network, making manipulation easier to produce and harder to pin down. Soon, financial markets may be full of artificial intelligences, each of which not only responds to events throughout the system but also exerts influence at each node, too. Existing theories of jurisdiction are already strained when, say, a hacker, server, internet service provider, and victim reside in different locations. Artificial intelligence now threatens to implicate hundreds of jurisdictions or none at all. Writing laws that hold individuals accountable in the abstract—or in absentia—may be a fascinating activity for an academic, but it would be unlikely to serve the public interest.

Specifically, artificial intelligence challenges the very notion of action by an individual. Instead, activity takes on a distributive nature, making it even more ephemeral. This is doubly true of artificial intelligence programs operating in the interconnected financial market.[147] The actions of one high-frequency trading algorithm resound across the market as other algorithms rapidly respond.[148] For example, the algorithm deployed by Navinder Singh Sarao, of the 2010 "Flash Crash," heightened market volatility not through the trades it issued alone, but instead by prompting a surge of trades by *other*

Fraud & Abuse Act, 18 U.S.C. § 1030 (1986) (main US federal law under which cyber criminals are prosecuted); United States v. Saavedra, 223 F.3d 85, 86 (2d Cir. 2000) ("[In] today's wired world of telecommunication and technology, it is often difficult to determine exactly where a crime was committed, since different elements may be widely scattered in both time and space, and those elements may not coincide with the accused's actual presence."); United States v. Muench, 694 F.2d 28, 33 (2d Cir. 1982) ("The intent to cause effects within the United States also makes it reasonable to apply to persons outside United States territory a statute which is not expressly extraterritorial in scope."); Roger A. Grimes, *Why It's So Hard to Prosecute Cyber Criminals*, CSO (Dec. 6, 2016), https://perma.cc/PMD3-N6B8; HARRY MATZ & SIVASHREE SUNDARAM, INTER-AM. DRUG ABUSE CONTROL COMM'N, DRUGS IN CYBERSPACE: UNDERSTANDING & INVESTIGATING DIVERSION & DISTRIBUTION OF CONTROLLED SUBSTANCES VIA THE INTERNET 18 (2006) ("The nature of the Internet is such that perpetrators may continue to run websites which dispense pharmaceuticals, regardless of prohibitive legislation, due to the anonymity involved in conducting such an operation."); Press Release, Dep't of Just., International Law Enforcement Operation Targeting Opioid Traffickers on the Darknet Results in over 170 Arrests Worldwide and the Seizure of Weapons, Drugs and over $6.5 Million (Sep. 22, 2020), https://perma.cc/TH6R-VN2P ("Riding the wave of technological advances, criminals attempt to further hide their activities within the dark web through virtual private networks and tails, presenting new challenges to law enforcement in the enduring battle against illegal drugs.").

147. *Cf.* SEC, DIV. OF ECON. & RISK ANALYSIS, U.S. CREDIT MARKETS: INTERCONNECTEDNESS AND THE EFFECTS OF THE COVID-19 ECONOMIC SHOCK 1 (2020) (describing how the economic shock brought by the Covid-19 pandemic revealed the interdependencies of financial markets, intermediaries, and participants).

148. *See generally* Yadav, *supra* note 11, at 1079 (explaining how the behavior of high-frequency trading algorithms potentiate the rapid compounding of errors across a financial market).

algorithms.[149] A decision by one artificial intelligence often produces outcomes that branch across the marketplace. As a result, where the conduct of one artificial intelligence ends and another's begins may prove a challenge to pin down.[150]

More important, the picture becomes infinitely more complicated when the actor "masking" the activity is an artificial intelligence. It is one thing to bemoan the fact that one cannot get ahold of an actor in another jurisdiction, but quite another thing when the "actor" making the choices—that is, the artificial intelligence—does not technically exist anywhere.

Finally, framing the question as "Where does the activity take place?" ignores the kernel of where it began. Is the law only interested in parties engaging in the activity? What about those who designed the activity or those who are facilitating it?[151] The law

149. See text accompanying notes 125-132 above for a discussion of the 2010 "Flash Crash" and its underlying causes.

150. *See* Yadav, *supra* note 11, at 1079 ("The trajectory of such systemic ripple effects can also be unexpected and difficult to predict.").

151. Copyright law offers a fruitful example of how legal regimes can distinguish between various stakeholders in the digital age, demonstrating a sensitivity between creators, facilitators and infringers. The Digital Millennium Copyright Act provides a safe harbor to internet service providers, absolving them from responsibility for the infringement that may occur via their network. *See* The Digital Millennium Copyright Act, 17 U.S.C. § 512; UMG Recordings, Inc. v. Shelter Capital Partners LLC, 718 F.3d 1006, 1036 (9th Cir. 2013) (holding that a video-sharing website qualified for safe-harbor protection); *see also* Jane C. Ginsburg & Luke Ali Budiardjo, *Liability for Providing Hyperlinks to Copyright-Infringing Content: International and Comparative Law Perspectives*, 41 COLUM. J.L. & ARTS 153, 200-208 (2018); John Blevins, *Uncertainty as Enforcement Mechanism: The New Expansion of Secondary Copyright Liability to Internet Platforms*, 34 CARDOZO L. REV. 1821, 1834-40 (2013). Not protected from liability, however, are platforms for users to post pirated works. *See* A&M Recs., Inc. v. Napster, Inc., 239 F.3d 1004, 1022-1024 (9th Cir. 2001), as amended (Apr. 3, 2001) (affirming that Napster, a database where users could post and access pirated music, was vicariously liable for copyright infringement); *cf.* Perfect 10, Inc. v. CCBill LLC, 488 F.3d 1102, 1114 (9th Cir. 2007) (holding that hosting password-hacking websites is not in and of itself proof of infringement). Protected platforms may also violate their immunity if they pass a "red flag" test demonstrating "subjective" or "objective" knowledge of infringement. *See* Viacom Int'l, Inc. v. YouTube, Inc., 676 F.3d 19, 31 (2d Cir. 2012) ("[T]he actual knowledge provision turns on whether the provider actually or 'subjectively' knew of specific infringement, while the red flag provision turns on whether the provider was *subjectively aware* of facts that would have made the specific infringement 'objectively' obvious to a reasonable person."); Ventura Content, Ltd. v. Motherless, Inc., 885 F.3d 597, 608-612 (9th Cir. 2018) (holding that plaintiff failed to establish that defendant knew infringing material was being uploaded to site and so defendant satisfied conditions for safe harbor). In a similar vein, the DMCA prohibits technology or services that are designed to infringe copy (e.g., descramblers), while protecting technology that may be used to violate copyright (e.g., CD re-writers). *See* Digitial Millenium Copyright Act, 17 U.S.C § 1201; *see generally The Digital Millennium Copyright Act*, 304 CORPORATE COUNSEL'S PRIMERS 1 (2019). Even protected technology (i.e., technology not designed to violate copyright), however, may be liable for copyright infringement if it is promoted for that purpose. *See* Metro-Goldwyn-Mayer Studios Inc. v. Grokster, Ltd., 545 U.S. 913, 919 (2005) ("We hold that one who distributes a device with the object of promoting its use to infringe copyright, as shown by clear expression or other affirmative steps taken to foster infringement, is liable for the resulting acts of infringement by third parties.").

must address each of these issues, as well as the resulting reverberations.

In sum, current legal regimes, designed with the governance of individuals and entities in mind, stand to be confused by the growth of artificial intelligence in the financial industry. Artificial intelligence lacks intent and is not intuitively located, factors that would challenge the prosecution of its actions under existing rules. Moreover, the past and current focus on size is increasingly inappropriate for artificially intelligent programs, which can empower rogue individuals to wreak outsized havoc on a market. The measures that have sought to regulate financial trading algorithms thus far are limited and reactive. The coming impact of artificial intelligence, however, demands a more comprehensive approach.

III. Touchpoint Regulation

As described in Part I, artificial intelligence in the financial industry enhances the capacity of what humans (or even ordinary predictive analytic programs) can accomplish. As artificial intelligence systems hypothesize, test, learn, and iterate—over and over again, all without human mediation—they travel far beyond the realm that any current legal system is prepared to manage. To solve the types of tantalizing questions raised as one explores the implications of artificial intelligence in the financial space, the law must develop an entirely new paradigm.

A. The Analogy of Flight

An intangible, such as a capacity or an activity, can be difficult to wrap one's mind around.[152] Nevertheless, the regulation of air travel in the United States provides a wonderful, albeit imperfect, analogy for regulating artificial intelligence in the financial industry. Consider the range of potential uses for air travel. Aviation can be used to transmit passengers, allowing humans to materialize 3,000 miles across the country in a matter of hours. It can deliver life-saving vaccines or cans of food. It can also be used to drop bombs on other nations, and it can be hijacked, so that the vehicles of flight crash into buildings, becoming bombs themselves. In short, aviation can be used by

152. For other proposals for regulating artificial intelligence in general or in the financial industry, see Stefan Heiss, *Towards Optimal Liability for Artificial Intelligence: Lessons from the European Union's Proposals of 2020*, 12 HASTINGS SCI. & TECH. L.J. 186 (2021) (advancing a risk-based approach for regulating AI); Bathaee, *supra* note 10 (proposing that the law should determine liability for AI according to "how much supervision and deference the AI receives, the training, validation, and test of the artificial intelligence, and the a priori constraints imposed on the artificial intelligence"); Mark A. Lemley & Bryan Casey, *Remedies for Robots*, 86 U. CHI. L. REV. 1311 (2019) (suggesting a harms-based approach for regulating AI); Lauren H. Scholz, *Algorithmic Contracts*, 20 STAN. TECH. L. REV. 128 (2017) (using examples from the high-speed trading context and recommending that AI should be treated as a constructive agent in the context of contracts created by AI); Azzutti, *supra* note 8 (discussing the so-called "human-in-the-loop" rule, which would improve the "traceability" and accountability of AI programs operating in capital markets by ensuring that humans are involved in all AI processes); *see also* Dirk A. Zetzsche, Douglas Arner, Ross Buckley & Brian W. Tang, *Artificial Intelligence in Finance: Putting the Human in the Loop* 19-20, CFTE Academic Paper Series No. 1, (2020), https://ssrn.com/abstract=3531711.

people—and entities—for good or for ill, part of waging war or thriving in peace.

Air travel is also subject to errors both human and mechanical, as well as to unpredictable disasters ranging from collisions with birds to freak weather patterns. The interactions and interests of various parties engaged in the activity of air travel must be intricately managed through air traffic control. Safety—both for those who travel and for those who might experience the externalities of noisy planes in their neighborhoods, refuse thrown from the skies, or plane parts crashing to the ground—must be safeguarded. Some individuals who engage in air travel (e.g., pilots) must be licensed, others must be subject to inspection (plane manufacturers), and still others must be identified (passengers).

Of course, air travel is not a perfect analogy. Current aviation does not employ artificial intelligence. A plane's autopilot will not choose to take a detour over forbidden airspace in North Korea because the flight path is much more direct and the likelihood of being shot down, while not nonexistent, is low.

Nevertheless, the extensive governance of the activity of flight is managed by controlling the points at which humans and entities come in contact with systems that the government can control, such as pilots, planes, airports, and air space, along with layering other controls on top of commercial actors who can be held accountable. This, we believe, provides a starting point for developing a paradigm for regulating artificial intelligence in the financial industry. We call the paradigm "touchpoint regulation."

B. Touchpoint Regulation Defined

Touchpoint regulation begins with the notion that certain aspects of U.S. financial markets operate as desirable entry points. Stock exchanges, commodity exchanges, and even banks provide entry points for those who wish to dabble in or dive into financial markets. These are the touchpoints that bring activities involving financial artificial intelligence into contact with spaces where governmental authority may hold sway. These are not the only pathways that could be used. In theory, exchanges could develop on the dark web, and alternative currencies, such as bitcoin and non-fungible tokens (NFTs),[153] already blossom. Nevertheless, even those using alternative currencies may wish to turn them into cold, hard cash at some point, and dark web exchanges are likely to be limited. In other words, as with planes, pilots could choose to take off from or land on a highway, but they probably wouldn't get very far.

Various regulations and requirements can be crafted that would allow the creative and exciting use of artificial intelligence while cabining its unethical and destructive potentials. These can range from the licensing of vendors of financial artificial

153. Non-fungible tokens are "one-of-a-kind" digital assets certifying ownership of a digital file (such as an image, a meme, or a social media post) that can be bought and sold. Records of NFT transactions are stored on the blockchain, a shared ledger maintained across thousands of computers that also records trades of cryptocurrency such as bitcoin. While tokenizing a digital file as an NFT does not prevent non-owners from making copies of the file (much as one can make copies of any other digital file), owning the NFT is intended to serve as proof that one owns the "original" work. As of August 2021, NFTs have sold for millions and even tens of millions of dollars. *See, e.g.*, *What Are NFTs and Why Are Some Worth Millions?*, BBC NEWS (Sep. 23, 2021), https://perma.cc/F7ZX-XWZ3.

intelligence to the identification of users of all sizes, to safety requirements for trading firms and exchanges, to updated and modernized mens rea requirements. We do not intend, with this article, to draft model legislation or define each regulatory nook and cranny. The goal, instead, is to establish the framework for a comprehensive legal structure. Specifically, we must begin thinking of our financial system as we think of commercial air travel, in which those who wish to step on board—or provide for or ferry those who do—are identifiable, identified, and subject to relevant rules of the road. Our current approach of engaging in limited, piecemeal parries in the hopes of avoiding the worst of the blows is unsustainable. We must think more broadly and boldly.

C. Frameworks for AI Governance

Within the general scaffolding provided above, we now turn to the ways in which the law should assign blame and responsibility. We begin from the perspective that the artificial intelligence itself makes the choices. Although such choices ostensibly are made within the bounds of the programming provided, programming can be quite general. The interpretations of those programming instructions and the pathways chosen can be startling to humans, if not downright unfathomable and unpredictable at times. Despite this potentially unknown and unknowable territory, the law cannot afford to simply abdicate its role, allowing society to be held hostage to the claim that "the algorithm made me do it."[154] Unless we are ready to hold the artificial intelligence accountable, the law must develop benchmarks for deciding when to assign culpability in the human realm.

Society enjoys many things that are useful but unpredictable in potentially dangerous ways, from dynamite to air travel. When a plane crashes because of ice on the wings or wind shear, society does not simply say, "Well, nature is unpredictable." Rather, the law regulates numerous points, from pilot training by the airline, to warning systems installed by the plane's manufacturer, to air traffic control systems by the government. In the context of air travel, each of these provides a touchpoint for regulation. The law needs an analogous approach for regulating the many touchpoints of artificial intelligence in the financial sector.

1. Three Types of Evil

In this context, we suggest that culpability can be organized according to three types of harms: the evil you planned, the evil you could have predicted, and the evil that was entirely unpredictable. In more stately legal terms, one could call these programmed harms, reasonably predictable harms, and unpredictable harms. Recall the *Coscia* case, in which programmers designed the algorithm to operate as a decoy for pumping the market and influencing other trading algorithms.[155] *Coscia* would be an example of a programmed harm. The level of liability should be at the highest when

154. *See supra* text accompanying notes 12-13.

155. United States v. Coscia, 866 F.3d 782, 789 (7th Cir. 2017).

the artificial intelligence is designed to produce the harm that resulted.

In contrast, an artificial intelligence might not have been directly programmed to produce the harm that resulted, but nevertheless, one could reasonably have predicted the harm. This level of liability would mirror a negligence standard.[156] For example, an artificial intelligence designed to trade in fractions of a second without fail-safes built in to halt its operations past a certain level of unusual trading could constitute reasonably predictable harm. One might not know what the harm will be and where the damage will occur, but when a plane without properly functioning warning lights crashes, the resulting harm can be characterized as reasonably predictable.

And of course, what is reasonably predictable to the programmer of an artificial intelligence may not be reasonably predictable to a stock trader or even a stock exchange. As we will describe in the Subpart below on different levels of responsibility for different touchpoint actors, the definition of what is reasonably predictable[157] may vary according to the actor.[158]

---

156. Negligence, generally, is the failure to act with a reasonable or ordinary level of care; as is pertinent to AI governance, the failure to act at all when there is a duty to act also constitutes negligence. *See generally* NEGLIGENCE, LEGAL INFO. INST., https://perma.cc/9JJT-NAXN (archived Aug. 5, 2021). A canonical test for assessing the liability of negligent actors in tort law is dubbed the "Hand Formula" after Judge Learned Hand's opinion. *See* United States v. Carroll Towing Co., 159 F.2d 169, 173 (2d Cir. 1947) (holding that if the probability of loss multiplied by the gravity of loss exceeds the burden of taking precautions, then the party with a duty to take precautions has liability to do so). In the context of artificial intelligence governance, for instance, a financial trading outfit (e.g., Knight Capital) might be held to introduce guardrails or other fail-safes for its algorithms if doing so is cheaper or less burdensome than would be the risk-adjusted damage caused by its malfunctioning in the marketplace. *See also supra* Part I.B. for a discussion of negligence alongside other liability standards in the financial industry. Some existing statutes accept negligence, without having to demonstrate scienter, as a basis for liability—if only in specific contexts such as IPOs and investment advising. A negligence standard for pleading artificial intelligence misconduct, as proposed here, would set an even lower burden of proof than a recklessness standard, which requires a more pronounced departure from the standards of ordinary care. *See* Fletcher, *supra* note 78, at 163 ("[I]f the algorithm's learning is the result of negligence, the recklessness standard would not be sufficient to hold the programmer liable.").

157. Courts have considerable experience establishing the boundaries of reasonableness in particular circumstances. *See, e.g.*, Brandon L. Garrett, *Constitutional Reasonableness*, 102 MINN. L. REV. 61, 70 (2017) ("The reasonable person standard is best known from negligence law, where courts have long used it to set an objective standard based on ordinary care in the relevant circumstances. In tort law, the flexibility of the concept of reasonable care may be a weakness, but also its strength, giving courts the ability (in theory, at least) 'to arrive at the correct judgment in a fact-dependent context,' even if the concept is 'frustratingly imprecise,' as Professor James Gibson puts it." (quoting James Gibson, *Doctrinal Feedback and (Un)Reasonable Care*, 94 VA. L. REV. 1641, 1643 (2008))); *see also* H.L.A. HART, THE CONCEPT OF LAW 128-29 (1961) (citing negligence's reasonable care standard as a pervasive example of courts' issuing *ex post facto* situational judgments to effect a "simple rule" where the "varied" "range of circumstances" renders *ex ante* lawmaking impossible).

158. For an example of how laws adjust expectations depending on the actor in question, see note 151 above, citing the different treatments of users and infringers under copyright law.

Artificial intelligence, however, continues to amaze, confound, delight, and frustrate even those who create it. Some harms, indeed, may be totally unpredictable. Recall the scenario in which an artificial intelligence, instructed to study and capitalize on social media trends to maximize profits, engages in extensive retweeting of false information.[159] Those involved truly might have had no idea that the artificial intelligence would develop the capacity to open a Twitter account and start retweeting.[160]

For cases in which the harm was truly unpredictable, one would expect the level of responsibility to be at the lowest. Of course, actors and entities might still be responsible for ensuring that the damage did not occur in the future. And at some point, the simple existence of errors becomes a predictable outcome in itself.[161]

2. Varieties of Actors

In the prior Subpart on the level of responsibility, we studiously avoided describing who to hold responsible for the various levels of harm. Turning to this topic, we suggest that the obligations—and resulting liabilities—of touchpoint actors should be organized according to whether they are users, intermediaries, or creators.

Consider the following example: A retail trader (we will call her "Grandma Heloise") manages to purchase an illegal artificial intelligence trading program that is designed to engage in stock manipulation and circumvent any warning triggers set by exchanges. The program's sales information clearly describes what it is designed to do, and Grandma Heloise thinks that is just fine, given the state of the world. The harms, in this case, are completely predictable to both Grandma Heloise and the programmer. Nevertheless, the law might choose to assign a higher level of punishment to the programmer, who manufactured the tool that made the harm possible, than to the individual user. Although one could argue that without the user, there is no market, the programmer, and the programmer's skills, may be a focal point of activities that can affect lots of trading and lots of traders. The law may wish to lean more heavily on those who facilitate what others merely dream about.[162] This approach models the experience of copyright law, in which creators of tools that facilitate pirating of copyrighted material are treated differently from those who are merely users, in terms of punishment.[163]

159. *See supra* text accompanying notes 133-135 (discussing the 2013 "Hash Crash").

160. *See id.*; *cf.* Bathaee, *supra* note 83, at 911-12 (posing a variant of this hypothetical in which an artificial intelligence creates false tweets).

161. *See, e.g.*, *supra* text accompanying notes 30-36 (giving examples of artificial intelligence mishaps).

162. To return to copyright law, for instance, Napster, a platform for sharing pirated music, was found indirectly liable for infringement of music copyright posted by its users. *See* A&M Recs., Inc. v. Napster, Inc., 239 F.3d 1004, 1022-1024 (9th Cir. 2001).

163. Under the Digital Millennium Copyright Act, circumventing copyright or trafficking in devices that enable circumvention (e.g., descramblers) is illegal, but not the use of those circumvented materials themselves. *See* 17 U.S.C. § 1201(a)(1)(A); *see also supra* note 151. *See generally* Robert C. Denicola, *Access Controls, Rights Protection, and Circumvention: Interpreting the Digital Millennium Copyright Act to Preserve Noninfringing Use*, 31 COLUM. J.L.

Treating different actors differently applies not just to levels of punishment but also to types of obligations. Following the copyright analogy above, a different set of obligations, safe harbors, and liabilities apply to those who are gatekeepers of the digital world, such as internet service providers and platform operators. In the same vein, the law might wish to assign a different set of obligations and liabilities to the exchanges and brokerage houses who serve intermediary roles. For example, some financial intermediaries are well-positioned to insist on licensing or identifying the parties who pass through them, in the same way that airport authorities check the identification of passengers. Other intermediaries, particularly exchanges, are well-positioned to identify harm as it unfolds. As a result, particular credentialing and identification obligations might best be assigned to these intermediaries. Exploring any of these issues, however, begins with accepting the paradigm that different financial market players will have different levels of obligation and liability. Following this pathway allows a more comprehensive framework in which all actors can bear some amount of obligation and responsibility. No one can simply say, "It's the artificial intelligence's fault."

The thorniest questions, however, revolve around the inability to identify and reach actors in the system. As legal doctrines from copyright to criminal law to international banking have discovered, a law on the books is of little use when one cannot reach the parties you wish to punish or on whom you wish to impose an obligation. With the distributive nature of artificial intelligence in finance, there may be circumstances in which we still may not know who you are just from grabbing ahold of actors through the financial system.

This could be mitigated, to some extent, by assigning the highest level of responsibility and obligation to unknown actors, creating incentives for identification. Moreover, the law could decide to place the burden of non-identification on those who interact with the unidentifiable. For example, as a user, you may ordinarily have little liability in comparison to the developer, but if you buy from an untraceable developer, the law could decide that you assume the programmer's liability. The system could also make anonymous use a crime. Both of these examples are analogous to the regulation of handguns, in that possession of an unregistered handgun constitutes a crime.[164] One could argue that certain types of artificial intelligence programs are properly analogous to loaded weapons. Finally, the need for profits to flow through the banking system provides other opportunities for identification and regulation.

The problem becomes somewhat more manageable with a system in which the

& ARTS 209 (2008). The statute also requires the Librarian of Congress to, every three years, publish a list of copyrighted works for which the anti-circumvention clause, because it harms non-infringing users of a material, no longer applies. This provision further distinguishes users from the active infringers. 17 U.S.C. § 1201(a)(1)(D); *see also, e.g.*, Maria Scheid, *2018 DMCA Section 1201 Exemptions Announced*, OHIO STATE UNIV. LIBRARIES (Mar. 20, 2019), https://perma.cc/5DXS-UNBL (describing the 2018 exemptions to the Digital Millennium Copyright Act, classes of works that included out-of-production computer programs for museum preservation and snippets of films for critical commentary). *See generally* David Nimmer, *A Riff on Fair Use in the Digital Millennium Copyright Act*, 148 U. PENN. L. REV. 673 (2000) (describing how and why the DMCA treats users and facilitators differently).

164. *See, e.g.*, N.Y. PENAL LAW §§ 265.01-04 (Consol. 2021).

overarching norm is for all players to be identified. And, of course, society could decide that no law is perfect and that it is content to leave some battles unfought.

Finally, the paradigms of categorization according to level of harm and categorization according to type of actor work well together in tandem. Recall that the law would assign a mid-range level of liability when one could have predicted the harm. What is reasonably predictable to the programmer of an artificial intelligence is different from what may be reasonably predictable to a stock trader or even a stock exchange. Thus, framing responsibility along the lines of different types of players and different types of harm provides the backbone for addressing a myriad of tricky issues regarding the use of artificial intelligence in the financial sector.

Conclusion

The artificial intelligence market is growing by leaps and bounds, with some projecting its revenue to increase tenfold between 2021 and 2028.[165] Its growth heralds an ever-deeper infusion of artificial intelligence into the financial industry—and, accordingly, an ever-greater need for legal and regulatory systems up to the task of accounting for the missteps and manipulations those processes will facilitate. Yet as situations like the Flash Crash and the Knight Capital trading glitch have already intimated, artificial intelligence poses a unique trial to frameworks predicated on the scienter of individuals and entities comprising groups of individuals. Lacking the capacity to act with intent, artificial intelligence can hardly be prosecuted, and if its actions are unprecedented, then neither does it make sense to blame its creator. Compounding the problem are the conditions that, first, even small players—once armed with artificially intelligent tools—can have a devastating impact on financial markets, and second, algorithms lack delimited places of residence.

Given these challenges, hewing to existing legal frameworks will prove more and more untenable in the years to come. However, just as policy makers once had to manage human beings' newfound capacity for flight, so, too, should financial regulators implement a comprehensive framework of touchpoint regulation, in which artificial intelligence would be held accountable at touchpoints where its use comes into contact with the broader financial system. Culpability could then be assessed depending on whether the harm caused was programmed, predictable, or unpredictable; and actors could be defined according to their status as users, intermediaries, or creators. Such a framework would have the flexibility and scope to account even for consequences that lie beyond the current boundaries of human imagination, while leaving room for artificial intelligence to reach its revolutionary potential. The future is bright if we navigate carefully.

165. GRAND VIEW RESEARCH, ARTIFICIAL INTELLIGENCE MARKET SIZE, SHARE & TRENDS ANALYSIS REPORT BY SOLUTION, BY TECHNOLOGY (DEEP LEARNING, MACHINE LEARNING, NATURAL LANGUAGE PROCESSING, MACHINE VISION), BY END USE, BY REGION, & SEGMENT FORECASTS, 2021-2028 (2021), https://perma.cc/GLM4-5CYR (archived Nov. 21, 2021).